

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:48:35 ; Search time 544 Seconds

(without alignments)
1615.956 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattttcttggttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------|
| 1 | 97.8 | 25.7 | 517 | 9 | US-11-096-568A-8355 |
| 2 | 70 | 18.4 | 876 | 9 | US-11-096-568A-29719 |
| 3 | 66.8 | 17.5 | 957 | 9 | US-11-087-099-102 |
| 4 | 66.2 | 17.4 | 896 | 7 | US-10-714-887-273 |
| 5 | 64 | 16.8 | 994 | 9 | US-11-096-568A-23894 |
| 6 | 63.2 | 16.6 | 1724 | 9 | US-11-096-568A-4492 |
| 7 | 55.2 | 14.5 | 866 | 7 | US-10-714-887-245 |
| 8 | 53.6 | 14.1 | 1254 | 9 | US-11-096-568A-21266 |
| 9 | 42.6 | 11.2 | 2891 | 8 | US-10-750-185-38981 |
| 10 | 42.6 | 11.2 | 2891 | 8 | US-10-750-623-38981 |
| 11 | 40.2 | 10.6 | 632 | 6 | US-09-925-065A-664600 |
| 12 | 38.8 | 10.2 | 527 | 6 | US-09-925-065A-525531 |
| 13 | 38.8 | 10.2 | 590 | 6 | US-09-925-065A-496715 |
| 14 | 38.8 | 10.2 | 607 | 6 | US-09-925-065A-496714 |
| 15 | 38.8 | 10.2 | 607 | 6 | US-09-925-065A-496716 |
| 16 | 38.8 | 10.2 | 667 | 6 | US-09-925-065A-944139 |
| 17 | 38.8 | 10.2 | 667 | 6 | US-09-925-065A-944140 |
| 18 | 38.6 | 10.1 | 605 | 6 | US-09-925-065A-580584 |
| 19 | 38.6 | 10.1 | 605 | 6 | US-09-925-065A-580585 |
| 20 | 38 | 10.0 | 759 | 6 | US-09-925-065A-925387 |

| | | | | | | |
|------|------|-----|---------|----|-----------------------|-------------------|
| C 21 | 37.8 | 9.9 | 562 | 6 | US-09-925-065A-416894 | Sequence 416894, |
| C 22 | 37.6 | 9.9 | 717 | 9 | US-11-096-568A-4827 | Sequence 4827, Ap |
| C 23 | 37.6 | 9.9 | 2208 | 8 | US-10-750-185-26159 | Sequence 26159, A |
| C 24 | 37.6 | 9.9 | 2208 | 8 | US-10-750-623-26159 | Sequence 26159, A |
| C 25 | 37.4 | 9.8 | 687411 | 7 | US-10-330-773-26 | Sequence 26, Appl |
| C 26 | 37.2 | 9.8 | 759 | 6 | US-09-925-065A-925386 | Sequence 925386, |
| C 27 | 37 | 9.7 | 599 | 6 | US-09-925-065A-387198 | Sequence 387198, |
| C 28 | 37 | 9.7 | 5455 | 8 | US-10-240-708-33 | Sequence 33, Appl |
| C 29 | 36.8 | 9.7 | 587 | 6 | US-09-925-065A-274001 | Sequence 274001, |
| C 30 | 36.8 | 9.7 | 588 | 6 | US-09-925-065A-561779 | Sequence 561779, |
| C 31 | 36.8 | 9.7 | 591 | 6 | US-09-925-065A-312318 | Sequence 312318, |
| C 32 | 36.8 | 9.7 | 634 | 6 | US-09-925-065A-487340 | Sequence 487340, |
| C 33 | 36.8 | 9.7 | 634 | 6 | US-09-925-065A-487341 | Sequence 487341, |
| C 34 | 36.6 | 9.6 | 616 | 6 | US-09-925-065A-694182 | Sequence 694182, |
| C 35 | 36.6 | 9.6 | 1128 | 12 | US-11-091-883-244 | Sequence 244, App |
| C 36 | 36.4 | 9.6 | 1135 | 8 | US-10-750-185-38542 | Sequence 38542, A |
| C 37 | 36.4 | 9.6 | 1135 | 8 | US-10-750-623-39542 | Sequence 39542, A |
| C 38 | 36.2 | 9.5 | 467 | 6 | US-09-925-065A-226581 | Sequence 226581, |
| C 39 | 36.2 | 9.5 | 615 | 6 | US-09-925-065A-730683 | Sequence 730683, |
| C 40 | 36 | 9.4 | 4453 | 12 | US-11-011-332A-99 | Sequence 99, Appl |
| C 41 | 36 | 9.4 | 1125000 | 8 | US-10-995-561-13286 | Sequence 13286, A |
| C 42 | 35.8 | 9.4 | 398 | 6 | US-09-925-065A-217343 | Sequence 217343, |
| C 43 | 35.8 | 9.4 | 398 | 6 | US-09-925-065A-217344 | Sequence 217344, |
| C 44 | 35.8 | 9.4 | 472 | 6 | US-09-925-065A-624778 | Sequence 624778, |
| C 45 | 35.6 | 9.3 | 523 | 6 | US-09-925-065A-201478 | Sequence 201478, |

ALIGNMENTS

RESULT 1

US-11-096-568A-8355

; Sequence 8355, Application US/11096568A

; Publication No. US20060048240A1

; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nickolai et al.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide

; TITLE OF INVENTION: Therby

; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A

; CURRENT FILING DATE: 2005-04-01

; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 8355

; LENGTH: 517

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(517)

; OTHER INFORMATION: Ceres Seq. ID no. 15225172

US-11-096-568A-8355

Query Match 25.7%; Score 97.8; DB 9; Length 517;
Best Local Similarity 69.8%; Pred. No. 6.1e-13;
Matches 132; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 34 | TGGACGTTTAGTCAAAAACAAGATGTTTCGAGAGGGCTTGGCAGTTTACGACAGGACACA | 93 |
| Db | 90 | TGGATCTCGAAGCAGAACAGAGATTTGAGATGCGCTTGCATCTTCGACAGGACACC | 149 |
| Qy | 94 | CCGACCGATGCGACAAATGTGGCAAAGCTGCGAGGGGAACTGTAGAAAGTGAAG | 153 |
| Db | 150 | CGAGCAGGTGCGACACCGTGGCCAGGCGCTGCGAGGAGGAAACCGTGGAGGAAGTGA | 209 |
| Qy | 154 | CGCCACTATGACATTCTCGTCGAGGATCTCATCAACATCGAGACTGGTGTGTCCTTTG | 213 |
| Db | 210 | AGGCATTATGAGAAGCTCGTGAAGATGTTGAGGAGATTGAGGAAGGTTCACGTGCCCTC | 269 |
| Qy | 214 | CCCAATTAC | 222 |
| Db | 270 | CCCAATTAC | 278 |

Db 222 AGATGTCATGAGCAATCTAAGCTTGAAGAAGACCTCTTCGATATCGAAGCACT 281
QY 204 TGTCCCTTTGCCCAATTAC 222
Db 282 TGTCCCGATCCCGGTTAC 300

RESULT 5
US-11-096-568A-23894
; Sequence 23894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23894
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(994)
; OTHER INFORMATION: Ceres Seq. ID no. 12416292
US-11-096-568A-23894

Query Match 16.8%; Score 64; DB 9; Length 994;
Best Local Similarity 60.2%; Pred. No. 3e-05;
Matches 106; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 32 CATGGAGTTTGTCAAAACAAGATGTTTCGAGAGGCCTTGGCAGTTTACGACAAGGACA 91
Db 202 CGTGGACGGCGGAGGAGAACAGCTGTTTCGAGAAGGCACTGGCGCAGATCGACCGGAAG 261
QY 92 CACCCGACGATGGCACAATGTGGCAAAAGCTGTCCGAGGGGAAACTGTAGAAGAGTGA 151
Db 262 CGCCGACAGTGGGAGAGTGGCGCGGTGCTGCCCTTGGAGACGGTTCGACGACGTGA 321
QY 152 AGCGCATATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGTGTCGTGTC 207
Db 322 GGAGCCAGTACCAACGCCCTCGAGAGGACGTGGCGGTCTATCGAGGCGCGCGGCTC 377

RESULT 6
US-11-096-568A-4492
; Sequence 4492, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4492
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1724)
; OTHER INFORMATION: Ceres Seq. ID no. 13638441
US-11-096-568A-4492

Query Match 16.6%; Score 63.2; DB 9; Length 1724;
Best Local Similarity 56.9%; Pred. No. 5.1e-05;
Matches 116; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 25 ATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGCCTTGGCAGTTTACGAC 84
Db 572 AGCACAATATGACCTCTGAAGAGAACAAGCTCTTTGAAATGCTCTTTCGAGTGCATGAT 631
QY 85 AAGGACACACCCGACCGCATGGCACAATGTGGCAAAAGCTGTTCGAGGGGAAAACTGTAGAA 144
Db 632 AAGGACACCCCGATCGGTGGCACAGATGGCTGAGATATTCCTGGAAAGACAGTGGTT 691
QY 145 GAAGTGAAGCGGCACCTATGACATTCCTCGTCGAGGATCTCATCAACATCGAGACTGGTCGT 204
Db 692 GATGTGATAAGGCAGTACAAAGGATTTGGAAGTAGATGTTAGCAATATAGAAGCTGGTTG 751
QY 205 GTCCCTTTGCCCAATTACAGACC 228
Db 752 ATTCCAGTTCTCGGCTATAGTAGC 775

RESULT 7
US-10-714-887-245
; Sequence 245, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 245
; LENGTH: 866
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2701 reference sequence; predicted polypeptide sequence is par
US-10-714-887-245

Query Match 14.5%; Score 55.2; DB 7; Length 866;
Best Local Similarity 55.9%; Pred. No. 0.0028;
Matches 105; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
QY 34 TGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGGCCTTGGCAGCTTTACGACAAGGACACA 93
|||
Db 142 TGGACTAAAGAGAGAAAGATGTTTCGAACGAGCTCTTTCGATATACGCTGAAGACTCG 201
|||
QY 94 CCCGACCGATGGCACAATGTGGCAAAAGCTGTCGGAGGGAAACTGTAGAGAAGTGAAG 153
|||
Db 202 CCTGATCGCTGGTAAAGTTGCTTCATGATCCCTGGAAAGACTGTTTTTGTATGTTATG 261
|||
QY 154 CGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTCGTGTGCCCTTTG 213
|||
Db 262 AACCAATATAGTAACTTGAAGAAGCGTTTCGATATTTGAACGAGGACGTGTTCCCAIT 321
|||
QY 214 CCCAATTA 221
|||
Db 322 CCTGGTTA 329
|||

RESULT 8
US-11-096-568A-21266
; Sequence 21266, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21266
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1254)
; OTHER INFORMATION: Ceres Seq. ID no. 12402516
US-11-096-568A-21266

Query Match 14.1%; Score 53.6; DB 9; Length 1254;
Best Local Similarity 59.08; Pred. No. 0.007;
Matches 92; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 82 GACAAGGACACACCCGACCGATGSCACAATGTGGCAAAAGCTGTGGAGGGAAACTGTGA 141
|||
Db 259 GAGGAGGACGGGACGCGAGGTGGGAGAAGCTAGCGGAGGCGGTGCGAGGGGAGACGCCG 318
|||
QY 142 GAAGAAGTGAAGCCCACTATGACATCTCGTCGAGGATCTCATCAACATCGAGACTGGT 201
|||
Db 319 GAGGAGGTGAGCGGCGCACTACGAGCTGTGTGGAGGACGTGCGACGCGCATCGAGTCGGC 378
|||
QY 202 CGTCTCCCTTGGCCCAATTACAAGACCTTTCGAATCT 237
|||
Db 379 CGCTCCGCTTCCGGGTACGCGGCTGACGGCGCT 414
|||

RESULT 9
US-10-750-185-38981
; Sequence 38981, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38981
; LENGTH: 2891
; TYPE: DNA
; ORGANISM: Bovine 19866881325079
US-10-750-185-38981

Query Match 11.2%; Score 42.6; DB 8; Length 2891;
Best Local Similarity 56.9%; Pred. No. 2.6;
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 239 ACTCAAGAAGCATCAATGACTTTGACACAGGATATATACTAAATATCTATATATGATGC 298
|||
Db 1995 ACTGAATATGTTTCATTTGTAGATGTAACAGGCAATTAACATAAGTCATATTTTCAGCCAG 2054
|||
QY 299 TCTCGATATATTTTGATAATCATTTAGTCATTTTGAGAAATCTCTCAAAAAGTTCTTG 358
|||
Db 2055 ACTCTAAATATTTCTAATAATCAATTTTATAAAATTTATCTTCTGAGTTGAAATTTT 2114
|||
QY 359 TAAGTTATATTTCTTTG 375
|||
Db 2115 TGTACTATATTAATCTGTG 2131
|||

RESULT 10
US-10-750-623-38981
; Sequence 38981, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38981
; LENGTH: 2891
; TYPE: DNA
; ORGANISM: Bovine 19866881325079
US-10-750-623-38981

Query Match 11.2%; Score 42.6; DB 8; Length 2891;
Best Local Similarity 56.9%; Pred. No. 2.6;
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 239 ACTCAAGAAGCATCAATGACTTTGACACAGGATATATACTAAATATCTATATATGATGC 298
|||
Db 1995 ACTGAATATGTTTCATTTGTAGATGTAACAGGCAATTAACATAAGTCATATTTTCAGCCAG 2054
|||
QY 299 TCTCGATATATTTTGATAATCATTTAGTCATTTTGAGAAATCTCTCAAAAAGTTCTTG 358
|||
Db 2055 ACTCTAAATATTTCTAATAATCAATTTTATAAAATTTATCTTCTGAGTTGAAATTTT 2114
|||
QY 359 TAAGTTATATTTCTTTG 375
|||
Db 2115 TGTACTATATTAATCTGTG 2131
|||

RESULT 11
US-09-925-065A-664600/c
; Sequence 664600, Application US/09925065A
; Publication No. US20040181048A1
```

```

Db      258 TAAAGAGCTATATATAAATATATATATAAATATATATAAATATATATAAATATATATAAATATATATA 139
Qy      311 TTGATAATCATTTCTAGTGATTTTCAGAAAATCTCTCAAAAAGTTCTCTGTAAGTTTATATTT 370
Db      198 TATATAAGCTGGGCAATTTCTAAGGAAAATTCAAACCAAAAAGTTGTGGCCAGTTAAATTA 139
Qy      371 CTTT 374
Db      138 TCTT 135

RESULT 13
US-09-925-065A-496715/c
; Sequence 496715, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496715
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496715

Query Match      10.2%; Score 38.8; DB 6; Length 590;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0

Qy      248 GCATCAATGACTTTGACACAGGTATATATACTAAATATCTATATATGCTCTCGATAT 307
Db      146 GCATATATTATCATGATACATAATTTATGTTATACAATATTTTATATCTTCATTTTCAGTAT 87
Qy      308 ATTTCGATAATCATTTCTAGTGATTTTCAGAAAATTTCTCAAAAAGTTCTTTGTAA 361
Db      86 AATTGGTTTGTCATCTTGTATTTATTTTATGCAATTTAAATAATATTATTGTGA 33

RESULT 14
US-09-925-065A-496714/c
; Sequence 496714, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:46:11 ; Search time 573 Seconds
(without alignments)
5498.487 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattctttgtgttaa 381

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------------------|--------------------|
| 1 | 381 | 100.0 | 381 | 8 | US-10-697-787-1 | Sequence 1, Appli |
| 2 | 381 | 100.0 | 381 | 9 | US-10-512-600-1 | Sequence 1, Appli |
| 3 | 120.6 | 31.7 | 760 | 8 | US-10-767-795-270 | Sequence 270, App |
| c 4 | 119 | 31.2 | 543 | 7 | US-10-021-323-16167 | Sequence 16167, A |
| 5 | 119 | 31.2 | 779 | 8 | US-10-767-795-271 | Sequence 271, App |
| 6 | 116.8 | 30.7 | 678 | 7 | US-10-424-599-6644 | Sequence 6644, Ap |
| 7 | 110.8 | 29.1 | 704 | 7 | US-10-425-114-14832 | Sequence 14832, A |
| 8 | 110.8 | 29.1 | 730 | 7 | US-10-424-599-135729 | Sequence 135729, A |
| c 9 | 110 | 28.9 | 466 | 7 | US-10-260-238-4786 | Sequence 4786, Ap |
| 10 | 108.4 | 28.5 | 577 | 7 | US-10-021-323-13836 | Sequence 13836, A |
| c 11 | 107.4 | 28.2 | 505 | 7 | US-10-021-323-16005 | Sequence 16005, A |
| 12 | 107.4 | 28.2 | 532 | 7 | US-10-021-323-14482 | Sequence 14482, A |
| 13 | 107.4 | 28.2 | 538 | 7 | US-10-021-323-11476 | Sequence 11476, A |
| c 14 | 107.4 | 28.2 | 542 | 7 | US-10-021-323-11563 | Sequence 11563, A |
| 15 | 107.4 | 28.2 | 604 | 8 | US-10-767-795-5000 | Sequence 5000, Ap |
| 16 | 106 | 27.8 | 679 | 7 | US-10-424-599-103390 | Sequence 103390, A |
| c 17 | 104.8 | 27.5 | 556 | 3 | US-09-770-152-473 | Sequence 473, App |
| 18 | 104.8 | 27.5 | 578 | 6 | US-10-225-066A-359 | Sequence 359, App |
| 19 | 104.8 | 27.5 | 578 | 7 | US-10-374-780A-2719 | Sequence 2719, Ap |
| 20 | 104.8 | 27.5 | 578 | 9 | US-10-225-066A-359 | Sequence 359, App |
| 21 | 103.8 | 27.2 | 636 | 8 | US-10-767-795-5934 | Sequence 5934, Ap |
| 22 | 103.4 | 27.1 | 584 | 7 | US-10-021-323-9298 | Sequence 9298, Ap |
| 23 | 98.6 | 25.9 | 507 | 7 | US-10-424-599-5904 | Sequence 5904, Ap |

ALIGNMENTS

RESULT 1

US-10-697-787-1
; Sequence 1, Application US/10697787
; Publication No. US20040216182A1
; GENERAL INFORMATION:
; APPLICANT: Agronomics, LLC
; TITLE OF INVENTION: Generation of Plants with Improved Pathogen Resistance and
; FILE REFERENCE: AG03-071C
; CURRENT APPLICATION NUMBER: US/10/697,787
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/US03/12981
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-697-787-1

| | | | | |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match | 100.0% | Score 381; | DB 8; | Length 381; |
| Best Local Similarity | 100.0% | Pred. No. 2.8e-94; | | |
| Matches 381; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTTTC | 60 | |
| Db | 1 | ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTTTC | 60 | |
| Qy | 61 | GAGAGGGCTTTGGCAGTTTACGACAAAGGACACCCGACCGATGGCAATGTGGCAAAA | 120 | |
| Db | 61 | GAGAGGGCTTTGGCAGTTTACGACAAAGGACACCCGACCGATGGCAATGTGGCAAAA | 120 | |
| Qy | 121 | GCTGTGGAGGAAACCTGTAGAAGTGAAGCGCCACTATGACATTCCTGTCGAGGAT | 180 | |
| Db | 121 | GCTGTGGAGGAAACCTGTAGAAGTGAAGCGCCACTATGACATTCCTGTCGAGGAT | 180 | |
| Qy | 181 | CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGAATCTTAC | 240 | |
| Db | 181 | CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGAATCTTAC | 240 | |
| Qy | 241 | TCAAGAAGCATCAATGACTTTTGACACAAAGGTATATACTATAATATCTATATGATGCTC | 300 | |
| Db | 241 | TCAAGAAGCATCAATGACTTTTGACACAAAGGTATATACTATAATATCTATATGATGCTC | 300 | |

QY 212 TGCCCAATTACAGACCTT 230
| | | | |
Db 325 TCCCGACTATTGGACCGT 307
| | | | |

RESULT 5

US-10-767-795-271
; Sequence 271, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; NUMBER OF SEQ ID NOS: 2004-01-30
; SEQ ID NO 271
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C42883_1
US-10-767-795-271

Query Match 31.2%; Score 119; DB 8; Length 779;
Best Local Similarity 74.9%; Pred. No. 4.3e-22;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 32 CATGGACGTTTGTAGTCAAAACAAGATGTTCCGAGAGGCGCTTGGCAGTTCAGCAAGGACA 91
| | | | |
Db 172 CATGGACAGCCAGCAAAACAAGATTTCCGAAGGCGTTAGCTGTTTACGAAGGACA 231
| | | | |
QY 92 CACCCGACGATGGCACAATGTGGCAAAAGCTGTCCGAGGGGAAAACTGTAGAGGAAGTGA 151
| | | | |
Db 232 CACCAGATCGTTGGTACAAATGTTGCTAAAGCTGTGGAGAGAAAACTGTTGAGGAAGTGA 291
| | | | |
QY 152 AGCGCCATGACATTCCTCGAGGATCTCATCAATCGAGACTGGTGTGTGTCCTT 211
| | | | |
Db 292 AGAAGCATATGAGCTTCTTCTTGAAGATGTTAGACGCATCGAGTCGGGTCGGGTTCCTT 351
| | | | |

QY 212 TGCCCAATTACAGACCTT 230
| | | | |
Db 352 TCCCGACTATTGGACCGT 370
| | | | |

RESULT 6

US-10-424-599-6644
; Sequence 6644, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6644
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106008C.1
US-10-424-599-6644

Query Match 30.7%; Score 116.8; DB 7; Length 678;

Best Local Similarity 70.1%; Pred. No. 1.6e-21;
Matches 157; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 21 TTCAATCTCACCATCGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTA 80
| | | | |
Db 238 TTCTGACTCTCTTGGACGCCCAAAAAGCAAGCTGTTTGAAGAAAGCACTTGCACAAATA 297
| | | | |
QY 81 CGACAAGGACACACCCGACCGATGACAAATGTGGCAAAAGCTGTCCGAGGGGAAAACTGT 140
| | | | |
Db 298 TGACAAGGATACCCCTGAGCGCTGGCAGATGTAGCCAANGCAGTAGTGGAAAAATCTGC 357
| | | | |
QY 141 AGAAGAAGTGAAGCGCCACTATGACATTTCTCGTGAGGATCTCATCAACATCGAGACTGG 200
| | | | |
Db 358 AGATGAAGTTAAGAGACACTATGAAATCTCTTGGAGGATCTCAGACACATAGAGTCTGG 417
| | | | |
QY 201 TGTGTCTCTTGGCCCAATTACAAAGACCTTCGAATCTTAACCTAA 244
| | | | |
Db 418 CCGTGTCTCTTCTCCCAAGTACAAGTCCACAGGAAGCAGCACCA 461
| | | | |

RESULT 7

US-10-425-114-14832
; Sequence 14832, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14832
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3028-039-A6_FLI
US-10-425-114-14832

Query Match 29.1%; Score 110.8; DB 7; Length 704;
Best Local Similarity 68.1%; Pred. No. 7.3e-20;
Matches 154; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGACGTTTAGTCAAAACAAGATGTTTC 60
| | | | |
Db 150 ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTGTAAAGACAACAGGCGCTTT 209
| | | | |
QY 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
| | | | |
Db 210 GAAAAGGCTTTAGCTGTTTATGACAAGGACACTCTCTGACCGTTGGTACAATGTTCTCAT 269
| | | | |
QY 121 GCTGTCCGAGGAAAACTGTAGAGAAGTGAAGCGCCACTATGACATTTCTGTGAGGAT 180
| | | | |
Db 270 GCTGTGTGGCAAAACTCCAGAGGAAGTGAAGAAAGCACTACGAACTCCTGTTTTCAGGAT 329
| | | | |
QY 181 CTCATCAACATCGAGACTGGTGTGTCCTTTTGGCCCAATTACAAGA 226
| | | | |
Db 330 GTTAAGCATATTGAGTCTGGACGTTGTCCTTCCCAAAATTACAAGA 375
| | | | |

RESULT 8

US-10-424-599-135729
; Sequence 135729, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K

```
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 135729
; LENGTH: 730
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_93572C.1
US-10-424-599-135729
```

```
Query Match 29.1%; Score 110.8; DB 7; Length 730;
Best Local Similarity 68.1%; Pred. No. 7.4e-20;
Matches 154; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTC 60
DB 147 ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTGTAAAGGACAACAAGGCTTT 206
QY 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
DB 207 GAAAAGGCTTTAGCTGTTTATGACAAGGACACTCTCGACCGTTGGTACAAATGTTGCTCAT 266
QY 121 GCTGTGGGAGGAAACTGTAGAGAGTGAAGGCGCACATCATCACATCTCTCGTCGAGGAT 180
DB 267 GCTGTGGTGGCAAACTCCAGAGGAGTGAAGGCGCACATCATCACATCTCTCGTCGAGGAT 326
QY 181 CTCATCAACATCAGACATGCTGTCGTCCTTTGCCCAATTACAAGA 226
DB 327 GTTAAGCATATTGAGTCTGACGCTGTCGCATTCGCCATTCCTCAAAATTACAAGA 372
```

```
RESULT 9
US-10-260-238-4786/c
; Sequence 4786, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4786
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Musa acuminata
US-10-260-238-4786
```

```
Query Match 28.9%; Score 110; DB 7; Length 466;
Best Local Similarity 67.4%; Pred. No. 9.7e-20;
Matches 155; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```
QY 25 ATCTCACCATGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTTTACGAC 84
DB 407 AGCTCGTGTGGACCGCGAAGCAGAACAAAGATGTTTCGAGAAAGCCCTCGCGGTGTACGAC 348
QY 85 AAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTGGAGGGGAAAACCTGTAGAA 144
DB 347 AAGGACACCGCCGACCGCTGGCACAACGTTGGCGCGCGCTCGGCGGCAAGTCAGCGGAA 288
QY 145 GAAAGTGAAGCGGCACCTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGGTCGT 204
DB 287 GAAAGTGAAGCGGCACCTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGGTCGT 204
QY 205 GTCCCTTTGCCAAATTTACAAGACCTTTCGAATCTTAATCTCAAGAAGCATCAA 254
DB 227 ATGCCTCGAGCCAATTTACCGCTCTCTCCGCGCCACAGGGGATGAAGTGTCTGA 178
```

```
RESULT 10
US-10-021-323-13836
; Sequence 13836, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 13836
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-015-Q1-K6-E2
US-10-021-323-13836
```

```
Query Match 28.5%; Score 108.4; DB 7; Length 577;
Best Local Similarity 68.0%; Pred. No. 3e-19;
Matches 151; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 9 AAATCAAGAAGTTCAATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGGC 68
DB 89 AATGTCATCGATTTCAATGTTGTTTCATGGACGACCAACAACAAGATTCGAAGGGC 148
QY 69 CTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTTCGG 128
DB 149 TTTAGCTGTTTACGACAAGGACACACACAGATCGTTGTTGTTGTTGTTGTTGTTGTTGTTG 208
QY 129 AGGGAACCTGTAGAAGAGTGAAGCGCCCACTATGACATTTCTCGTCGAGACTCATCAA 188
DB 209 AGAGAAAACCTGTTTGAAGAGTGAAGAGCACTATGAGCTTCTTCTTGAAGATGTTAGACG 268
QY 189 CATCGAGACTGGTTCGTCCTTTGCCCAATTACAAGACCTT 230
DB 269 CATCGAGTCGGTTCGGGTTCTTTCCCGGACTATTGACCGT 310
```

```
RESULT 11
US-10-021-323-16005/c
; Sequence 16005, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 16005
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-N6-A10
US-10-021-323-16005

Query Match 28.2%; Score 107.4; DB 7; Length 505;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTAC 81
DB 466 TCAGGTTGCAACTGGAGCGCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 407
QY 82 GACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTTCGAGGGGAAAACCTGTA 141
DB 406 GACAAGGATACACCAAGCCGTTGGCACAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 347
QY 142 GAAGAGTGAAGGCCCACTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGTT 201
DB 346 GAGGAAGTGAAGTTGCATTACCAAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 287
QY 202 CGTGCCCTTTGCCCAATTACAAGA 226
DB 286 CACGTGCCTTTGCCCTTTACAAGA 262

RESULT 12
US-10-021-323-14482
; Sequence 14482, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14482
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-A10
US-10-021-323-14482

Query Match 28.2%; Score 107.4; DB 7; Length 532;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTAC 81
DB 56 TCAGGTTGCAACTGGAGCGCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 115
QY 82 GACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTTCGAGGGGAAAACCTGTA 141
DB 116 GACAAGGATACACCAAGCCGTTGGCACAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 175

QY 142 GAAGAGTGAAGGCCCACTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGTT 201
DB 176 GAGGAAGTGAAGTTGCATTACCAAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 235
QY 202 CGTGCCCTTTGCCCAATTACAAGA 226
DB 236 CACGTGCCTTTGCCCTTTACAAGA 260

RESULT 13
US-10-021-323-11476
; Sequence 11476, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 11476
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-002-Q1-K6-E10
US-10-021-323-11476

Query Match 28.2%; Score 107.4; DB 7; Length 538;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTAC 81
DB 33 TCAGGTTGCAACTGGAGCGCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 92
QY 82 GACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTTCGAGGGGAAAACCTGTA 141
DB 93 GACAAGGATACACCAAGCCGTTGGCACAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 152
QY 142 GAAGAGTGAAGGCCCACTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGTT 201
DB 153 GAGGAAGTGAAGTTGCATTACCAAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 212
QY 202 CGTGCCCTTTGCCCAATTACAAGA 226
DB 213 CACGTGCCTTTGCCCTTTACAAGA 237

RESULT 14
US-10-021-323-11563/c
; Sequence 11563, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880

Job time : 574 secs

```
; SEQ ID NO 11563
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-002-Q1-N6-E10
US-10-021-323-11563

Query Match      28.2%; Score 107.4; DB 7; Length 542;
Best Local Similarity 70.2%; Pred. No. 5.5e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy  22 TCAATCTCACCATTGGAGCTTTAGTCAAAAACAAGATGTTGAGAGGGCCTTGGCAGTTTAC 81
Db  506 TCAGGTTGGAACCTGGAGCCCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 447

Qy  82 GACAAGGACACACCCGACCGATGGCAATGTGGCAAAAGCTGTCGGAGGGAAACTGTA 141
Db  446 GACAAGGATACACAGACCGGTGGCAACAGCTAGCCAGGCTGTTGGGGGAAAGACCGTG 387

Qy  142 GAAGAAGTGAAGCGCCACTATGACATTTCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
Db  386 GAGGAAGTGAAGTTGCATTACCAAGACCTTTGTGGATGACATCAAGCAGATAGAGTCTGGG 327

Qy  202 CGTGTCCTTTGGCCCAATTACAAGA 226
Db  326 CACGTGCCCTTTGCCCCCTTACAAGA 302

RESULT 15
US-10-767-795-5000
; Sequence 5000, Application US/10767795
; Publication No. US2004018130A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 5000
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C53491_1
US-10-767-795-5000

Query Match      28.2%; Score 107.4; DB 8; Length 604;
Best Local Similarity 70.2%; Pred. No. 5.8e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy  22 TCAATCTCACCATTGGAGCTTTAGTCAAAAACAAGATGTTGAGAGGGCCTTGGCAGTTTAC 81
Db  99 TCAGGTTGGAACCTGGAGCCCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 158

Qy  82 GACAAGGACACACCCGACCGATGGCAATGTGGCAAAAGCTGTCGGAGGGAAACTGTA 141
Db  159 GACAAGGATACACAGACCGGTGGCAACAGCTAGCCAGGCTGTTGGGGGAAAGACCGTG 218

Qy  142 GAAGAAGTGAAGCGCCACTATGACATTTCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
Db  219 GAGGAAGTGAAGTTGCATTACCAAGACCTTTGTGGATGACATCAAGCAGATAGAGTCTGGG 278

Qy  202 CGTGTCCTTTGGCCCAATTACAAGA 226
Db  279 CACGTGCCCTTTGCCCCCTTACAAGA 303

Search completed: March 10, 2006, 20:03:22
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:29:10 ; Search time 2275 Seconds

(without alignments)

9519.716 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_cov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_scs.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|-------------|
| 1 | 381 | 100.0 | 381 | 15 | AY519525 | Arabidops |
| 2 | 381 | 100.0 | 68041 | 15 | AC025814 | Arabidops |
| 3 | 273.4 | 71.8 | 294 | 6 | CS138004 | Sequence |
| 4 | 151.4 | 39.7 | 303 | 15 | AY519524 | Arabidops |
| 5 | 151.4 | 39.7 | 334 | 15 | BT005657 | Arabidops |
| 6 | 151.4 | 39.7 | 580 | 15 | AK119034 | Arabidops |
| 7 | 149.4 | 39.2 | 120977 | 15 | AC025808 | Genomic s |
| 8 | 108.8 | 28.6 | 612 | 15 | LES277944 | Lycopersi |
| 9 | 108.8 | 28.6 | 3605 | 15 | AJ583670 | Lycopersi |
| 10 | 108.4 | 28.5 | 246 | 15 | BT011255 | Arabidops |
| 11 | 108.4 | 28.5 | 506 | 15 | BT010770 | Arabidops |
| 12 | 108.4 | 28.5 | 100108 | 14 | AP008111 | Lotus cor |
| 13 | 106.6 | 28.0 | 97268 | 15 | AP004546 | Lotus cor |
| 14 | 106.6 | 28.0 | 102550 | 14 | AP008039 | Lotus cor |
| 15 | 106.6 | 28.0 | 108878 | 14 | AP007808 | Lotus cor |
| 16 | 104.8 | 27.5 | 306 | 15 | AY519526 | Arabidops |
| 17 | 104.8 | 27.5 | 337 | 15 | BT008554 | Arabidops |
| 18 | 104.8 | 27.5 | 570 | 15 | BT008698 | Arabidops |

| | | | | | | |
|----|-------|------|--------|----|--------------|-------------------|
| 19 | 104.8 | 27.5 | 578 | 6 | CS137904 | Sequence |
| 20 | 104.8 | 27.5 | 54433 | 15 | AC007119 | Arabidops |
| 21 | 103.2 | 27.1 | 3213 | 15 | AY954971 | Antirrhin |
| 22 | 102.8 | 27.0 | 108844 | 15 | AC157502 | Medicago |
| 23 | 99.4 | 26.1 | 512 | 15 | AK112054 | Oryza sat |
| 24 | 99 | 26.0 | 198354 | 15 | ATAP22 | Arabidops |
| 25 | 99 | 26.0 | 198750 | 15 | ATCHRIV85 | Arabidops |
| 26 | 96.2 | 25.2 | 34070 | 15 | AP007251 | Oryza sat |
| 27 | 96.2 | 25.2 | 110000 | 15 | AP008207_254 | Continuation (255 |
| 28 | 96.2 | 25.2 | 110220 | 14 | AP003688 | Oryza sat |
| 29 | 96.2 | 25.2 | 162297 | 15 | AP004317 | Oryza sat |
| 30 | 92.4 | 24.3 | 513 | 6 | CS137972 | Sequence |
| 31 | 92.4 | 24.3 | 92861 | 15 | AC006439 | Arabidops |
| 32 | 92 | 24.1 | 294 | 6 | CS137908 | Sequence |
| 33 | 92 | 24.1 | 294 | 15 | AY519527 | Arabidops |
| 34 | 92 | 24.1 | 89035 | 15 | AT122F8 | Arabidops |
| 35 | 92 | 24.1 | 198987 | 15 | ATCHRIV90 | Arabidops |
| 36 | 91.6 | 24.0 | 91322 | 14 | AP007301 | Arabidops |
| 37 | 89.4 | 23.5 | 166907 | 14 | AC146974 | Lotus cor |
| 38 | 84.6 | 22.2 | 99587 | 15 | AC007858 | Zea mays |
| 39 | 84.6 | 22.2 | 110000 | 15 | AP008211_286 | Oryza sat |
| 40 | 84.6 | 22.2 | 156643 | 15 | AC120988 | Continuation (287 |
| 41 | 84.4 | 22.2 | 108564 | 14 | AC162787 | Oryza sat |
| 42 | 79.6 | 20.9 | 867 | 15 | AY519533 | Medicago |
| 43 | 79.6 | 20.9 | 898 | 15 | BT005473 | Arabidops |
| 44 | 79.6 | 20.9 | 1640 | 15 | AK118891 | Arabidops |
| 45 | 79.6 | 20.9 | 77758 | 15 | AB016885 | Arabidops |

ALIGNMENTS

| | | | | | | |
|-----------|--|---|--------|------|--------|-----------------|
| RESULT 1 | AY519525 | Arabidopsis thaliana MYB transcription factor (At1g75250) mRNA, complete cds. | 381 bp | mRNA | linear | PLN 07-FEB-2004 |
| LOCUS | AY519525 | Arabidopsis thaliana MYB transcription factor (At1g75250) mRNA, complete cds. | | | | |
| ACCESSION | AY519525.1 | GI:41618977 | | | | |
| VERSION | AY519525.1 | GI:41618977 | | | | |
| KEYWORDS | Arabidopsis thaliana (thale cress) | | | | | |
| SOURCE | Arabidopsis thaliana | | | | | |
| ORGANISM | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi | | | | | |
| REFERENCE | 1 (bases 1 to 381) | | | | | |
| AUTHORS | Qu,L. and Gu,H. | | | | | |
| TITLE | The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis | | | | | |
| JOURNAL | Unpublished | | | | | |
| AUTHORS | Qu,L. and Gu,H. | | | | | |
| TITLE | Direct Submission | | | | | |
| JOURNAL | Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China | | | | | |
| FEATURES | Location/Qualifiers | | | | | |
| source | 1..381 | | | | | |
| gene | /organism="Arabidopsis thaliana" | | | | | |
| CDS | /mol_type="mRNA" | | | | | |
| | /db_xref="taxon:3702" | | | | | |
| | 1..381 | | | | | |
| | /locus_tag="At1g75250" | | | | | |
| | 1..381 | | | | | |
| | /locus_tag="At1g75250" | | | | | |
| | /codon_start=1 | | | | | |
| | /product="MYB transcription factor" | | | | | |
| | /protein_id="AA09995.1" | | | | | |
| | /db_xref="GI:41618978" | | | | | |
| | /translation="MASNSRSSISPTFTSQNMFERALAVYDKDTPDRMHNVAKAVGG KTVSEVKRHVDILVDLNIETGRVPLPNYKTFPSNGRSINDFDTRYITKLYLMLSI YFDNHSSDFEKFQSKVLVYSISLV" | | | | | |
| ORIGIN | | | | | | |

[illegible]

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

```
FEATURES
  source
    1..334
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /db_xref="taxon:3702"
      /chromosome="1"
      /clone="U51233"
      /ecotype="Columbia"
      /note="This clone is in pUNI 51."
  gene
    1..334
      /gene="Atlg19510"
  CDS
    1..303
      /gene="Atlg19510"
      /codon_start=1
      /evidence=experimental
      /product="putative myb family transcription factor"
      /protein_id="AA064077.1"
      /db_xref="GI:28973505"
      /translation="MASSSMSSSSWTSKONKMPERALAAYDKDTPDRWQNVAKAVGS
      KSAEVEKSHYDLVLEDLMNIEQDLVPLPKYKTVGVGSKSRGIDDFDLRLMKMNRQ"
    3'UTR
      304..334
        /gene="Atlg19510"

ORIGIN
  Query Match      39.7%; Score 151.4; DB 15; Length 334;
  Best Local Similarity 73.2%; Pred. No. 5.5e-27;
  Matches 210; Conservative 0; Mismatches 71; Indels 6; Gaps 1;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
   |||||
DB 1 ATGGCTCTAGTTCTATGAGCTCGAGCTCTTCTTGAGCTCTAAGCAAAACAAGATGTTTC 60
   |||||

QY 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
   |||||
DB 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCTAAA 120
   |||||

QY 121 GCTGTCGAGGAGAACTGTAGAGAGTGAGCGCCACTATGACATTCCTCGCAGGAT 180
   |||||
DB 121 GCAGTTGGAAGTAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 180
   |||||

QY 181 CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGA-----A 234
   |||||
DB 181 CTCATGAACATCGAACAAGACTTAGTACCTTTGCCCTAATAACAACCGTCGATGTGA 240
   |||||

QY 235 TCTAACTCAAGAAGCATCAATCACTTTGACACAAGGTATATACTAA 281
   |||||
DB 241 AGTAAATCTAGAGGCATCGATGATTTTGATTTGAGGTAAATGAAGAA 287
   |||||

RESULT 6
AK119034
LOCUS
DEFINITION
  Arabidopsis thaliana Atlg19510 mRNA for putative myb-related
  protein, complete cds, clone: RAF21-36-F10.
ACCESSION
  AK119034
VERSION
  AK119034.1 GI:26453067
KEYWORDS
  FLI CDNA; CAP trapper.
SOURCE
  Arabidopsis thaliana
  Arabidopsis thaliana (thale cress)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1
    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
    Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
    Kawai, J., Hayashizaki, Y. and Shinozaki, K.
  Arabidopsis thaliana full-length cDNA
  Published Only in Database (2002)
  2 (bases 1 to 580)
    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
    Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,

FEATURES
  Location/Qualifiers
    1..580
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /db_xref="taxon:3702"
      /chromosome="1"
      /clone="RAF21-36-F10"
      /ecotype="Columbia"
      /note="common name: thale cress"
    1..580
      /gene="Atlg19510/F18014_31"
    88..390
      /gene="Atlg19510/F18014_31"
      /codon_start=1
      /product="putative myb-related protein"
      /protein_id="BAC43610.1"
      /db_xref="GI:26453068"
      /translation="MASSSMSSSSWTSKONKMPERALAAYDKDTPDRWQNVAKAVGS
      KSAEVEKSHYDLVLEDLMNIEQDLVPLPKYKTVGVGSKSRGIDDFDLRLMKMNRQ"

ORIGIN
  Query Match      39.7%; Score 151.4; DB 15; Length 580;
  Best Local Similarity 73.2%; Pred. No. 5.5e-27;
  Matches 210; Conservative 0; Mismatches 71; Indels 6; Gaps 1;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
   |||||
DB 88 ATGGCTCTAGTTCTATGAGCTCGAGCTCTTCTTGAGCTCTAAGCAAAACAAGATGTTTC 147
   |||||

QY 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
   |||||
DB 148 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCTAAA 207
   |||||

QY 121 GCTGTCGAGGAGAACTGTAGAGAGTGAGCGCCACTATGACATTCCTCGCAGGAT 180
   |||||
DB 208 GCAGTTGGAAGTAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 267
   |||||

QY 181 CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGA-----A 234
   |||||
DB 268 CTCATGAACATCGAACAAGACTTAGTACCTTTGCCCTAATAACAACCGTCGATGTGA 327
   |||||

QY 235 TCTAACTCAAGAAGCATCAATCACTTTGACACAAGGTATATACTAA 281
   |||||
DB 328 AGTAAATCTAGAGGCATCGATGATTTTGATTTGAGGTAAATGAAGAA 374
   |||||

RESULT 7
AC025808/c
LOCUS
DEFINITION
  Arabidopsis thaliana AC025808
  genomic sequence for Arabidopsis thaliana BAC F18014 from
  chromosome I, complete sequence.
ACCESSION
  AC025808
VERSION
  AC025808.8 GI:7636235
KEYWORDS
  HTG.
SOURCE
  Arabidopsis thaliana (thale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1
    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
    Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
    Kawai, J., Hayashizaki, Y. and Shinozaki, K.
  Arabidopsis thaliana full-length cDNA
  Published Only in Database (2002)
  2 (bases 1 to 580)
    Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
    Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
```

Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail:msaki@gs.riken.go.jp,
URL: <http://pfgweb.gsc.riken.go.jp>, Tel: 81-45-503-9625,
Fax: 81-45-503-9586)
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified pBluescript vector.
Please visit our web site (<http://pfgweb.gsc.riken.go.jp/>) for
further details.

FEATURES

source

1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF21-36-F10"
/ecotype="Columbia"
/note="common name: thale cress"
1..580
/gene="Atlg19510/F18014_31"
88..390
/gene="Atlg19510/F18014_31"
/codon_start=1
/product="putative myb-related protein"
/protein_id="BAC43610.1"
/db_xref="GI:26453068"
/translation="MASSSMSSSSWTSKONKMPERALAAYDKDTPDRWQNVAKAVGS
KSAEVEKSHYDLVLEDLMNIEQDLVPLPKYKTVGVGSKSRGIDDFDLRLMKMNRQ"

gene

CDS

ORIGIN

Query Match 39.7%; Score 151.4; DB 15; Length 580;
Best Local Similarity 73.2%; Pred. No. 5.5e-27;
Matches 210; Conservative 0; Mismatches 71; Indels 6; Gaps 1;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
 |||||
DB 88 ATGGCTCTAGTTCTATGAGCTCGAGCTCTTCTTGAGCTCTAAGCAAAACAAGATGTTTC 147
 |||||

QY 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
 |||||
DB 148 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCACAAATGTGCTAAA 207
 |||||

QY 121 GCTGTCGAGGAGAACTGTAGAGAGTGAGCGCCACTATGACATTCCTCGCAGGAT 180
 |||||
DB 208 GCAGTTGGAAGTAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 267
 |||||

QY 181 CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGA-----A 234
 |||||
DB 268 CTCATGAACATCGAACAAGACTTAGTACCTTTGCCCTAATAACAACCGTCGATGTGA 327
 |||||

QY 235 TCTAACTCAAGAAGCATCAATCACTTTGACACAAGGTATATACTAA 281
 |||||
DB 328 AGTAAATCTAGAGGCATCGATGATTTTGATTTGAGGTAAATGAAGAA 374
 |||||

RESULT 7

AC025808/c
LOCUS
DEFINITION
 Arabidopsis thaliana AC025808
 genomic sequence for Arabidopsis thaliana BAC F18014 from
 chromosome I, complete sequence.
ACCESSION
 AC025808
VERSION
 AC025808.8 GI:7636235
KEYWORDS
 HTG.
SOURCE
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
 1
 Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
 Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
 Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Arabidopsis thaliana full-length cDNA
 Published Only in Database (2002)
 2 (bases 1 to 580)
 Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
 Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1
AUTHORS Sobolev,I., Chelnitsky,I., Barg,R. and Salts,Y.
TITLE The tomato early fruit specific gene LefSM1 defines a novel class of plant-specific SANT/MYB domain proteins1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3605)
AUTHORS Salts,Y
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2003) Salts Y., Plant Genetics, The Volcani Center, ARO, PO Box 6, Bat-Dagan, 50250, ISRAEL

FEATURES
source
1..3605
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
/chromosome="10"
repeat_region
1..930
/transposon="retrotransposon like sequence"
repeat_region
1232..1523
/transposon="foldback transposon TAP1R1"
gene
2385..3605
/gene="fem1"
TATA_signal
2385..2389
/gene="fem1"
mRNA
join(2429..2789,3355..3605)
/gene="fem1"
exon
2429..2789
/gene="fem1"
5'UTR
2429..2505
/gene="fem1"
CDS
2506..2772
/gene="fem1"
/codon_start=1
/product="SANT/MYB protein"
/protein_id="CAE47523.1"
/db_xref="GI:36783452"
intron
2790..3354
/gene="fem1"
exon
3355..3605
/gene="fem1"
polyA_signal
3586..3591
/gene="fem1"
polyA_site
3605
/gene="fem1"

ORIGIN
Query Match 28.6%; Score 108.8; DB 15; Length 3605;
Best Local Similarity 70.2%; Pred. No. 2.2e-16;
Matches 146; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy 20 GTTCAATCTCACCATGGACGTTTGTAGTCAAAACAGATGTTTCGAGAGGCGCTTGGCAGTTT 79
Db 2531 GTTCTCTGGATCATGGACTGCGAAACAAACAGGCGTTTGAGAGGCGATTGGCTGTGT 2590
Qy 80 ACACAAGGACACACCGACCGATGGCAACATGTGGCAAAAGCTGTGCGAGGGGAAACTG 139
Db 2591 ATGACAGGAACTCGGATGTTGTGTCATGTTGCTTAAGCAGTTGGAGGCAACACG 2650
Qy 140 TAGAAGAAGTGAAGCGCCATATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTG 199
Db 2651 CAGAAGAAGTGAAGAAGACACTATGAAATCCTTTTGGAGGATGTTTCTTTATTGACAATG 2710
Qy 200 GTCGTGTCCTTTCGCCAATTACAGAC 227

Db 2711 GTATGTCGCCCTTCCCAATAACAAAAC 2738

RESULT 10
BT011255
LOCUS Arabidopsis thaliana At4g36570 gene, complete cds. PLN 14-JAN-2004
DEFINITION Arabidopsis thaliana At4g36570 gene, complete cds.
ACCESSION BT011255
VERSION BT011255.1 GI:40823587
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 246)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 246)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2004) Salk Institute Genomic Analysis Laboratory (SIGNat), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
Location/Qualifiers
1..246
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="C62918"
/ecotype="Columbia"
/note="This clone is in pUNI 51"
CDS
1..246
/note="unknown protein"
/codon_start=1
/product="At4g36570"
/protein_id="AAR92291.1"
/db_xref="GI:40823588"
/translation="MASNSMSSSSWTRKENKLFERALATYDQDTDRWHNVARVGG KSAEVRHYELLIRDVNDIESGRIPHPYRSNNH"

ORIGIN
Query Match 28.5%; Score 108.4; DB 15; Length 246;
Best Local Similarity 68.0%; Pred. No. 2.8e-16;
Matches 151; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 1 ATGGCGTCAAACTCAAGAAGTTCATCTCACCATGACGCTTTAGTCAAAACAAGATGTTTC 60
Db 1 ATGGCTTCCAACTCAATGAGCTCTAGCGCTTCTTGACACACGTAAGGAGACAATAATTATTT 60
Qy 61 GAGAGGGCCTTGGCAGTTTACGACAGGACACACCGACCGCATGTCGACCAATGTGGCAAAA 120
Db 61 GAAAGGCGGTGGCTACATATGACCGAGGACCTCTCGACCGTTGGCATACGTTGCAGA 120
Qy 121 GCTGTGCGAGGGAACCTGTAGAAGAAGTGAAGCGCCACTATGACATTCCTGTCGAGGAT 180
Db 121 GCGGTTGCGGCAAAATCAGCTGAAGAAGTAAGCGACACTACGAGCTCCTCATAGGAT 180
Qy 181 CTCATCAACATCGAGACTGCTGTGTGTCCTTTTGCCCAATTAC 222
Db 181 GTCAATGACATTGAGTCAGGCGGCTTATCCACATCCCAATTAC 222

RESULT 11
BT010770
LOCUS Arabidopsis thaliana At4g36570 mRNA, complete cds. PLN 29-NOV-2003
DEFINITION Arabidopsis thaliana At4g36570 mRNA, complete cds.
ACCESSION BT010770
VERSION BT010770.1 GI:38566493
KEYWORDS FLI_CDNA.

SOURCE ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 506)
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.
TITLE Arabidopsis cDNA clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 506)
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2003) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

FEATURES source
1..506
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="S62918"
/ecotype="Columbia"
/note="This clone is in pUNI 51"
34..279
/note="unknown protein"
/codon_start=1
/product="At4g36570"
/protein_id="AA24137.1"
/db_xref="GI:38566494"
/translation="MASNSMSSASWTRKENKLFERLATYDQTPDRWNVARAVGG
KSAEVRHYELLIRDVNDIESGRYPHENYRNGNH"

CDS
5202: contig of 5202 bp in length
5302: gap of unknown length
5303: contig of 1264 bp in length
5566: gap of unknown length
5667: gap of unknown length
7760: contig of 1094 bp in length
7860: gap of unknown length
8624: contig of 764 bp in length
8625: gap of unknown length
8725: contig of 883 bp in length
9608: gap of unknown length
9708: contig of 1834 bp in length
11541: contig of unknown length
11641: gap of unknown length
11642: contig of 2154 bp in length
13796: gap of unknown length
13895: contig of 1847 bp in length
13896: gap of unknown length
15742: gap of unknown length
15843: contig of 3189 bp in length
19032: contig of unknown length
19132: contig of 2679 bp in length
21811: gap of unknown length
21911: contig of 2757 bp in length
24668: gap of unknown length
24768: contig of 5539 bp in length
30307: gap of unknown length
30406: gap of unknown length
36585: contig of 6179 bp in length
36586: gap of unknown length
42389: contig of 5704 bp in length
42390: gap of unknown length
44889: gap of unknown length
50891: contig of 8402 bp in length
50892: gap of unknown length
60141: contig of 9150 bp in length
60142: gap of unknown length
60241: gap of unknown length
60242: contig of 8837 bp in length
69079: gap of unknown length
69179: contig of 10310 bp in length
79489: gap of unknown length
79588: contig of 18889 bp in length
98478: gap of unknown length
98578: contig of 1531 bp in length.
Location/Qualifiers
1..100108
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/var_tety="japonicus"
/db_xref="taxon:34305"
/clone="LjT05D15"
/clone_lib="LjT library"
/note="TAC clone:TM1490, synonym:Lotus japonicus"
5203..5302
/estimated_length=unknown
5567..6666
/estimated_length=unknown
7761..7860
/estimated_length=unknown
8625..8724

FEATURES source
1..100108
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/var_tety="japonicus"
/db_xref="taxon:34305"
/clone="LjT05D15"
/clone_lib="LjT library"
/note="TAC clone:TM1490, synonym:Lotus japonicus"
5203..5302
/estimated_length=unknown
5567..6666
/estimated_length=unknown
7761..7860
/estimated_length=unknown
8625..8724

Query Match 28.5%; Score 108.4; DB 15; Length 506;
Best Local Similarity 68.0%; Pred. No. 2.8e-16;
Matches 151; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 1 ATGGCGTCAACTCAAGAGTTCAATCTCACATCGACGCTTAGTCAAAACAGATGTC 60
DB 34 ATGGCTTCAACTCAATGAGCTCTAGCGCTCTTGGACACGTAAGGAGAACAAATTATT 93
QY 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCGCCGATGGCACAATGTGCAAAA 120
DB 94 GAAAGGGCTTGGCTACATATGACCAGACACTCTGTGCGTTGGCATAACGTTCAAGA 153
QY 121 GCTGTCGAGGAAACCTGTAGAGAGTGAAGGCCCACTATGACATTTCTCGTCAGGAT 180
DB 154 GCCGTGTCGGGCAATCAGCTGAGAGAGTGAAGGCGACACTACGAGCTCTCATTAGGAT 213
QY 181 CTCATCAATCGACACTGCTGCTGCTCTTGGCCCAATTAC 222
DB 214 GTCAATGACATGATGATGAGGCGGTTATCCACATCCCAATTAC 255

RESULT 12
AP008111
LOCUS AP008111 100108 bp DNA linear HTG 28-DEC-2004
DEFINITION Lotus corniculatus var. japonicus clone LjT05D15, *** SEQUENCING IN
PROGRESS ***, 20 unordered pieces.
ACCESSION AP008111
VERSION AP008111.1 GI:56806418
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE 1
AUTHORS Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. XI. Sequence

FEATURES and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 100108)
Sato, S.
Direct Submission
Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 5202: contig of 5202 bp in length
* 5302: gap of unknown length
* 5303: contig of 1264 bp in length
* 5566: gap of unknown length
* 5667: gap of unknown length
* 7760: contig of 1094 bp in length
* 7860: gap of unknown length
* 8624: contig of 764 bp in length
* 8625: gap of unknown length
* 8725: contig of 883 bp in length
* 9608: gap of unknown length
* 9708: contig of 1834 bp in length
* 11541: contig of unknown length
* 11641: gap of unknown length
* 11642: contig of 2154 bp in length
* 13796: gap of unknown length
* 13895: contig of 1847 bp in length
* 13896: gap of unknown length
* 15742: gap of unknown length
* 15843: contig of 3189 bp in length
* 19032: contig of unknown length
* 19132: contig of 2679 bp in length
* 21811: gap of unknown length
* 21911: contig of 2757 bp in length
* 24668: gap of unknown length
* 24768: contig of 5539 bp in length
* 30307: gap of unknown length
* 30406: gap of unknown length
* 36585: contig of 6179 bp in length
* 36586: gap of unknown length
* 42389: contig of 5704 bp in length
* 42390: gap of unknown length
* 44889: gap of unknown length
* 50891: contig of 8402 bp in length
* 50892: gap of unknown length
* 60141: contig of 9150 bp in length
* 60142: gap of unknown length
* 60241: gap of unknown length
* 60242: contig of 8837 bp in length
* 69079: gap of unknown length
* 69179: contig of 10310 bp in length
* 79489: gap of unknown length
* 79588: contig of 18889 bp in length
* 98478: gap of unknown length
* 98578: contig of 1531 bp in length.
Location/Qualifiers
1..100108
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/var_tety="japonicus"
/db_xref="taxon:34305"
/clone="LjT05D15"
/clone_lib="LjT library"
/note="TAC clone:TM1490, synonym:Lotus japonicus"
5203..5302
/estimated_length=unknown
5567..6666
/estimated_length=unknown
7761..7860
/estimated_length=unknown
8625..8724

```
/estimated_length=unknown
9608. .9707
/estimated_length=unknown
11542. .11541
/estimated_length=unknown
13796. .13895
/estimated_length=unknown
15743. .15842
/estimated_length=unknown
19032. .19131
/estimated_length=unknown
21811. .21910
/estimated_length=unknown
24668. .24767
/estimated_length=unknown
30307. .30406
/estimated_length=unknown
36586. .36685
/estimated_length=unknown
42390. .42489
/estimated_length=unknown
50892. .50991
/estimated_length=unknown
60142. .60241
/estimated_length=unknown
69079. .69178
/estimated_length=unknown
79489. .79588
/estimated_length=unknown
98478. .98577
/estimated_length=unknown

ORIGIN
Query Match      28.5%; Score 108.4; DB 14; Length 100108;
Best Local Similarity 58.5%; Pred. No. 2.8e-16;
Matches 207; Conservative 0; Mismatches 146; Indels 1; Gaps 1;

QY 1 ATGGCGTCAAACTCAGAAGTTCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTTTC 60
Db 74473 ATGGCATCAAGCTCAATGTCATCTATGCTCATGGAATGCTAAGGACAATAAGCCCTT 74532
QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGCGCAAAA 120
Db 74533 GAAAGGCGCTAGCTGTTTATGACAGGAAACCCCGACAGATGGTACAAATGTTGCTCAA 74592
QY 121 GCTGTGCGAGGAAACTGTAGAGAAAGTGAAGCGCCACTATGACATTTCTCGTGGAGAT 180
Db 74593 GCTGTTGGTGGGAAGACCCAGAGGAAGTGAAGGCGCACTATGAGCTTCTCTTAGGAT 74652
QY 181 CTCATCAACATCGAGATGTCGTGTCCTCTTTCGCCAATTACAGACCTTCGAATCTTAAC 240
Db 74653 GTTAGGCACATTGAATCAGGCGCATGTTCCATTCCCAAACTACAAGACACTGGA-GGGTC 74711
QY 241 TCAAGAGCATCAATGACTTTGACACAGGATATATACTAAATATCTATATATGATGCTC 300
Db 74712 AGACAGGAGAAAAGGTAGTGCACATAAACTGTATCTTCATATTCCTCACTTGAACAAA 74771
QY 301 TCGATATATTTGATAATCATCTTCAGTGGATTTTGAAGAAATCTCTCAAAAAGTT 354
Db 74772 TGCMAATATGCAATCTTCACTTCTGCTGCTGCTGAAGTTGCTATATAGTT 74825

RESULT 13
AP004546/c
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
clone:LjT43N05, TM0178, complete sequence.
ACCESSION
AP004546
VERSION
AP004546.1 GI:17736913
KEYWORDS
HTG.
ORGANISM
Lotus corniculatus var. japonicus (Lotus japonicus)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

/estimated_length=unknown
9608. .9707
/estimated_length=unknown
11542. .11541
/estimated_length=unknown
13796. .13895
/estimated_length=unknown
15743. .15842
/estimated_length=unknown
19032. .19131
/estimated_length=unknown
21811. .21910
/estimated_length=unknown
24668. .24767
/estimated_length=unknown
30307. .30406
/estimated_length=unknown
36586. .36685
/estimated_length=unknown
42390. .42489
/estimated_length=unknown
50892. .50991
/estimated_length=unknown
60142. .60241
/estimated_length=unknown
69079. .69178
/estimated_length=unknown
79489. .79588
/estimated_length=unknown
98478. .98577
/estimated_length=unknown

ORIGIN
Query Match      28.0%; Score 106.6; DB 15; Length 97268;
Best Local Similarity 70.6%; Pred. No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 28 TCACCATGGACGCTTTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAG 87
Db 32507 TCTACTTGGAGTCCAAAACAGAACAGATGTTTGGAGAGGACCTTCGCAAAATATGACGAG 32448
QY 88 GACACACCCGACGATGGCACAAATGTGGCAAAAGCTGTGCGAGGGAAAACTGTAGAAGAA 147
Db 32447 GACACCCCTGAGCGGTGGCAAAATGTGCCAAAGCAGTTGGTGGCAAACTCAGTGGAGGAA 32388
QY 148 GTGAAGCGGCATATGACATTTCTCGTGGAGGATCTCATCAACATCGAGCTGTCGTGTC 207
Db 32387 GTCAAGAGACATATGAGAGACTTTGGAGGATCTTCAACACATTCAGTCTGCCCATGTT 32328
QY 208 CCTTTGCCCAATTACAAGACC 228
Db 32327 CCTTTCCCACTACAAATCC 32307

RESULT 14
AP008039/c
LOCUS
DEFINITION
Lotus corniculatus var. japonicus chromosome 1 clone LjT10P24, ***
SEQUENCING IN PROGRESS ***, 24 unordered pieces.
ACCESSION
AP008039
VERSION
AP008039.1 GI:56806347
KEYWORDS
HTG; HTGS, PHASE1.
SOURCE
Lotus corniculatus var. japonicus (Lotus japonicus)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE
1 Kaneko,T., Asamizu,E., Nakamura,Y., Sato S. and Tabata S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
REFERENCE 2 (bases 1 to 102550)
AUTHORS Sato,S.
TITLE Direct Submission
```

JOURNAL

Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 3089: contig of 3089 bp in length
* 3090 3189: gap of unknown length
* 3190 4482: contig of 1293 bp in length
* 4483 4582: gap of unknown length
* 4583 5591: contig of 1009 bp in length
* 5592 5691: gap of unknown length
* 5692 6980: contig of 1289 bp in length
* 6981 7080: gap of unknown length
* 7081 8737: contig of 1657 bp in length
* 8738 8837: gap of unknown length
* 8838 10414: contig of 1577 bp in length
* 10415 10514: gap of unknown length
* 10515 11793: contig of 1279 bp in length
* 11794 11893: gap of unknown length
* 11894 14747: contig of 2854 bp in length
* 14748 14847: gap of unknown length
* 14848 17118: contig of 2271 bp in length
* 17119 17218: gap of unknown length
* 17219 19795: contig of 2577 bp in length
* 19796 19895: gap of unknown length
* 19896 23048: contig of 3153 bp in length
* 23049 23148: gap of unknown length
* 23149 25124: contig of 1976 bp in length
* 25125 25157: gap of 33 bp
* 25158 25546: contig of 389 bp in length
* 25547 25646: gap of unknown length
* 25647 28751: contig of 3105 bp in length
* 28752 28851: gap of unknown length
* 28852 32078: contig of 3227 bp in length
* 32079 32178: gap of unknown length
* 32179 37013: contig of 4835 bp in length
* 37014 37113: gap of unknown length
* 37114 41223: contig of 4110 bp in length
* 41224 41324: gap of unknown length
* 41324 46917: contig of 5494 bp in length
* 46918 46917: gap of unknown length
* 46918 54494: contig of 7577 bp in length
* 54495 54594: gap of unknown length
* 54595 64935: contig of 10341 bp in length
* 64936 65035: gap of unknown length
* 65036 74018: contig of 8983 bp in length
* 74019 74118: gap of unknown length
* 74119 88264: contig of 14146 bp in length
* 88265 88364: gap of unknown length
* 88365 100720: contig of 12356 bp in length
* 100721 100820: gap of unknown length
* 100821 102550: contig of 1730 bp in length.
Location/Qualifiers
1..102550
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone="LjT10P24"
/clone_lib="LjT library"
/note="TAC clone:IM1388, synonym:Lotus japonicus"
3090..3189
/estimated_length=unknown
4483..4582

FEATURES
source

1..102550

/organism="Lotus corniculatus var. japonicus"

/mol_type="genomic DNA"

/variety="japonicus"

/db_xref="taxon:34305"

/chromosome="1"

/clone="LjT10P24"

/clone_lib="LjT library"

/note="TAC clone:IM1388, synonym:Lotus japonicus"

3090..3189

/estimated_length=unknown

4483..4582

gap

gap

gap /estimated_length=unknown
5592..5691
/estimated_length=unknown
5981..7080
/estimated_length=unknown
8738..8837
/estimated_length=unknown
10415..10514
/estimated_length=unknown
11794..11893
/estimated_length=unknown
14748..14847
/estimated_length=unknown
17119..17218
/estimated_length=unknown
19796..19895
/estimated_length=unknown
23049..23148
/estimated_length=unknown
25125..25157
/estimated_length=33
25547..25546
/estimated_length=unknown
28752..28851
/estimated_length=unknown
32079..32178
/estimated_length=unknown
37014..37113
/estimated_length=unknown
41224..41323
/estimated_length=unknown
46818..46917
/estimated_length=unknown
54495..54594
/estimated_length=unknown
64936..65035
/estimated_length=unknown
74019..74118
/estimated_length=unknown
88265..88364
/estimated_length=unknown
100721..100820
/estimated_length=unknown
ORIGIN

Query Match 28.0%; Score 106.6; DB 14; Length 102550;
Best Local Similarity 70.6%; Pred. No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 28 TCACCATGGACGTTTAGTCAAAACAAGATGTCGAGAGGGCTTCGACACAG 87
Db 96874 TCTACTTGGAGTCCAAACAAGACAAAGTGTGAGAGAGACATTCGACACAG 96815
QY 88 GACACACCCGACCGATGGCAACAATGTGCAAAAGCTCGAGAGGAAACTGTAGAAGAA 147
Db 96814 GACACCCCTGAGCGGTGGCAAAATGTGTCCAAGACAGTTGGTGGCAAAATCATGTGAGGAA 96755
QY 148 GTGAAGCCCACTATGACATTCCTCGTCGAGGATTCATCAACATCGAGACTGGTCTGTGTC 207
Db 96754 GTCAAGAGACACTATGAGAGACTCTTGGAGGATCTTCAAAACACATTCGAGTCTGCGCATGTT 96695
QY 208 CCTTTGCCCAATTACAGACC 228
Db 96694 CCCTTTCCCACTACAAATCC 96674

RESULT 15

LOCUS

DEFINITION

AP007808

AP007808

AP007808

AP007808

AP007808 108878 bp DNA linear HTG 28-DEC-2004
Lotus corniculatus var. japonicus chromosome 1 clone LjB04J07, ***
SEQUENCING IN PROGRESS ***, 24 unordered pieces.
AP007808
AP007808.1 GI:56806118

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS PHASE1.
Lotus corniculatus var. japonicus (Lotus japonicus)
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 108878)
Sato, S.
Direct Submission
Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp, Kisarazu,
URL: <http://www.kazusa.or.jp/>, Tel: 81-438-52-3935(ex.2337),
Fax: 81-438-52-3934)

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4606: contig of 4606 bp in length
* 4607 4706: gap of unknown length
* 4707 5430: contig of 744 bp in length
* 5451 5550: gap of unknown length
* 5551 6707: contig of 1157 bp in length
* 6708 6807: gap of unknown length
* 6808 8058: contig of 1251 bp in length
* 8059 8158: gap of unknown length
* 8159 9271: contig of 1113 bp in length
* 9272 9372: gap of unknown length
* 9372 10504: contig of 1132 bp in length
* 10504 12079: gap of unknown length
* 12080 12179: contig of 1476 bp in length
* 12180 13219: gap of unknown length
* 13220 15621: contig of 2302 bp in length
* 15622 15721: gap of unknown length
* 15722 17539: contig of 1818 bp in length
* 17540 17639: gap of unknown length
* 17640 19483: contig of 1843 bp in length
* 19483 20910: gap of unknown length
* 20911 23211: contig of 1328 bp in length
* 23212 23312: contig of 2201 bp in length
* 23312 26479: gap of unknown length
* 26480 26579: contig of 3168 bp in length
* 26580 29860: gap of unknown length
* 29861 29960: contig of 3281 bp in length
* 29961 36032: gap of unknown length
* 36033 36132: contig of 6072 bp in length
* 36133 39620: gap of unknown length
* 39621 39720: contig of 3488 bp in length
* 39721 45217: gap of unknown length
* 45218 45317: contig of 5497 bp in length
* 45318 49895: gap of unknown length
* 49896 58773: contig of 4578 bp in length
* 58774 58873: gap of unknown length
* 58874 67840: contig of 8778 bp in length
* 67841 79594: gap of unknown length
* 79595 93084: contig of 11754 bp in length
* 93084 93084: contig of 13390 bp in length

FEATURES
source

* 93085 93184: gap of unknown length
* 93185 108878: contig of 15694 bp in length.
1. 108878
Location/Qualifiers
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variant="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone_lib="LjB library"
/note="BAC clone: BM1105, synonym: Lotus japonicus"
4607..4706
/estimated_length=unknown
5451..5550
/estimated_length=unknown
6708..6807
/estimated_length=unknown
8059..8158
/estimated_length=unknown
9272..9371
/estimated_length=unknown
10504..10603
/estimated_length=unknown
12080..12179
/estimated_length=unknown
13220..13319
/estimated_length=unknown
15622..15721
/estimated_length=unknown
17540..17639
/estimated_length=unknown
19483..19582
/estimated_length=unknown
20911..21010
/estimated_length=unknown
23212..23311
/estimated_length=unknown
26480..26579
/estimated_length=unknown
29861..29960
/estimated_length=unknown
36033..36132
/estimated_length=unknown
39621..39720
/estimated_length=unknown
45218..45317
/estimated_length=unknown
49896..49995
/estimated_length=unknown
58774..58873
/estimated_length=unknown
67741..67840
/estimated_length=unknown
79595..79694
/estimated_length=unknown
93085..93184
/estimated_length=unknown

ORIGIN

Query Match 28.0%; Score 106.6; DB 14; Length 108878;
Best Local Similarity 70.8%; Pred.No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 28 TCACCATGGACGCTTTAGTCAAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAG 87
58515 TCTACTTGGAGTCCAAAACAGACAAAGTTGTTTGAGAGAGCACTTGCAAAATATGACGAG 58456
Db 88 GACACACCGACCGATGGCACAATGTGGCAAAAGCTGTCCGAGGGGAAAACCTGTAGAAGAA 147
58455 GACACCCCTGAGCGGTGGCAAAATGTGTCCAAAGCAGTTGGTGGCAATCAGTGGAGGAA 58396
QY 148 GTGAAGCGCCACTATGACATTCTCGTCGAGGACTCTCATCAACATCGAGACTGGTCGTGTC 207

Db 58395 GTCAAGAGACACTATGAGAGACTCTTGGAGGATCTCAAACACACATTGAGTCTGCGCATGTT 58336
Qy 208 CCTTGGCCCAATTACAGACC 228
Db 58335 CCTTTCCCACTACAAATCC 58315

Search completed: March 10, 2006, 19:26:15
Job time : 2280 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:36:58 ; Search time 136 Seconds
(without alignments)
4979.787 Million cell updates/sec

Title: US-10-697-787-1
Perfect score: 381
Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/6C_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/6D_COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfile61.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| C 1 | 69.6 | 18.3 | 574 | US-09-640-211A-1534 | Sequence 1534, Ap |
| 2 | 58.4 | 15.3 | 357 | US-09-640-211A-1581 | Sequence 1581, Ap |
| 3 | 50.4 | 13.2 | 383 | US-09-640-211A-1310 | Sequence 1310, Ap |
| 4 | 40.2 | 10.6 | 100863 | US-09-949-016-17031 | Sequence 17031, A |
| C 5 | 38.8 | 10.2 | 399 | US-09-621-976-8976 | Sequence 8976, Ap |
| 6 | 37.4 | 9.8 | 89892 | US-09-949-016-13667 | Sequence 13667, A |
| C 7 | 37.4 | 9.8 | 670689 | US-09-949-016-12505 | Sequence 12505, A |
| C 8 | 37.4 | 9.8 | 670690 | US-09-949-016-14207 | Sequence 14207, A |
| C 9 | 37 | 9.7 | 1141 | US-09-806-708B-22 | Sequence 22, Appl |
| 10 | 36.8 | 9.7 | 1689 | US-08-447-500-11 | Sequence 11, Appl |
| 11 | 36.8 | 9.7 | 1689 | US-08-453-866-11 | Sequence 11, Appl |
| 12 | 36.8 | 9.7 | 2385 | US-08-454-097-9 | Sequence 9, Appl |
| 13 | 36.8 | 9.7 | 2385 | US-08-468-036-41 | Sequence 41, Appl |
| 14 | 36.8 | 9.7 | 2385 | US-08-376-843-41 | Sequence 41, Appl |
| 15 | 36.8 | 9.7 | 2385 | US-08-185-359-9 | Sequence 9, Appl |
| 16 | 36.8 | 9.7 | 58356 | US-09-949-016-15563 | Sequence 15563, A |
| 17 | 36.6 | 9.6 | 601 | US-09-949-016-103973 | Sequence 103973, A |
| 18 | 36.6 | 9.6 | 8868 | US-08-815-809-7 | Sequence 7, Appl |
| C 19 | 36.6 | 9.6 | 9870 | US-09-245-928A-15 | Sequence 15, Appl |
| 20 | 36.6 | 9.6 | 19877 | US-08-816-155B-8 | Sequence 8, Appl |
| 21 | 36.6 | 9.6 | 19877 | US-09-079-587-8 | Sequence 8, Appl |
| 22 | 36.6 | 9.6 | 84839 | US-09-949-016-15816 | Sequence 15816, A |
| 23 | 36.6 | 9.6 | 147382 | US-09-949-016-14624 | Sequence 14624, A |
| C 24 | 36.4 | 9.6 | 16573 | US-09-949-016-14876 | Sequence 14876, A |

| | | | | | | |
|------|------|-----|--------|---|---------------------|-------------------|
| C 25 | 36 | 9.4 | 601 | 3 | US-09-949-016-78876 | Sequence 78876, A |
| C 26 | 36 | 9.4 | 784019 | 3 | US-09-949-016-14033 | Sequence 14033, A |
| C 27 | 36 | 9.4 | 828152 | 3 | US-09-949-016-12777 | Sequence 12777, A |
| 28 | 35.6 | 9.3 | 429 | 3 | US-09-621-976-9543 | Sequence 9543, Ap |
| 29 | 35.6 | 9.3 | 1860 | 3 | US-09-919-497-50 | Sequence 50, Appl |
| 30 | 35.6 | 9.3 | 2069 | 3 | US-09-949-016-5801 | Sequence 5801, Ap |
| 31 | 35.4 | 9.3 | 332 | 3 | US-09-640-211A-1824 | Sequence 1824, Ap |
| 32 | 35.2 | 9.2 | 32392 | 3 | US-09-662-254B-27 | Sequence 27, Appl |
| C 33 | 34.8 | 9.1 | 1230 | 3 | US-09-140-466-1 | Sequence 1, Appli |
| 34 | 34.6 | 9.1 | 1784 | 3 | US-09-486-581-1 | Sequence 1, Appli |
| 35 | 34.6 | 9.1 | 1784 | 3 | US-10-213-797-1 | Sequence 1, Appli |
| 36 | 34.6 | 9.1 | 38346 | 3 | US-09-949-016-14502 | Sequence 14502, A |
| 37 | 34.6 | 9.1 | 39318 | 3 | US-09-949-016-13798 | Sequence 13798, A |
| C 38 | 34.6 | 9.1 | 43117 | 3 | US-09-949-016-17589 | Sequence 17589, A |
| 39 | 34.6 | 9.1 | 52199 | 3 | US-09-949-016-12763 | Sequence 12763, A |
| 40 | 34.6 | 9.1 | 52203 | 3 | US-09-949-016-16330 | Sequence 16330, A |
| 41 | 34.4 | 9.0 | 290 | 3 | US-10-131-827-8636 | Sequence 8636, Ap |
| 42 | 34.4 | 9.0 | 601 | 3 | US-09-949-002-8929 | Sequence 8929, Ap |
| C 43 | 34.4 | 9.0 | 1885 | 3 | US-09-662-254B-41 | Sequence 41, Appl |
| 44 | 34.4 | 9.0 | 50000 | 3 | US-09-662-254B-24 | Sequence 24, Appl |
| C 45 | 34.4 | 9.0 | 80411 | 3 | US-09-949-016-15777 | Sequence 15777, A |

ALIGNMENTS

RESULT 1
US-09-640-211A-1534/C
; Sequence 1534, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1534
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1534

| | | | | |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match | 18.3% | Score 69.6; | DB 3; | Length 574; |
| Best Local Similarity | 59.7% | Pred. No. 2.3e-10; | | |
| Matches 117; | Conservative 0; | Mismatches 79; | Indels 0; | Gaps 0; |
| QY | 28 | TCACCATGGAGTTTGTAGTCAAAACAAAGATGTTTCGAGAGGGCTTGGCAGATTTCACGACAAG | 87 | |
| Db | 431 | TCCTCGTGGAGTAAGAAACAGGCAAGCGGTTTCGAGAATGCCCTGGCTACCCACATTCGAG | 372 | |
| QY | 88 | GACACACCCGACCATGTCACAAATGTCGCAAAAGCTGCGAGGGAACACTGTAGAGAA | 147 | |
| Db | 371 | GTTTCCCGGATGTTGGTGGAGAAGGTACGCTCCGACGTCGCCGGAACACCTAGAGGAG | 312 | |
| QY | 148 | GTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGGTGGTGTGTC | 207 | |
| Db | 311 | ATTAAGCTCCACTACGAGGACCTGGTGAAGATGTCATAGATAGAGGCTGGCCCGCTG | 252 | |
| QY | 208 | CCTTTGCCCAATTACA | 223 | |
| Db | 251 | CCTCTCGCGTCTCTACA | 236 | |

RESULT 2
US-09-640-211A-1581
; Sequence 1581, Application US/09640211A
; Patent No. 6833446

2

```
Db 277 RNSYRRAMRGSKWGGSGYRMYAGYRSSRWSYSAMWRKDKKQKWKGRSSWGRSTGY 218
QY 109 AATGTGGCAAAAGCTGCGAGGAGAAACGTAGAGAAAGTGAAGCGCACTATGACATT 168
Db 217 YAMWYKSWCTSRWYKWKRRKWKCTUSTKRTCYRSTYKWKAYTYKKRKKWTRWT 158
QY 169 CTGCTCGAGGATCTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACC 228
Db 157 YYYKYSMSKKTWRMKTAYVYTKRWK-MTRTKWTCTWCKCTTYMAGTWMYRYRYW 99
QY 229 TTGCAATCTAAGTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAAATCTA 288
Db 98 YYAKRAKWSKRCTWSTTCYCMKYMAKKCWSYMWNSMGMKWSMMWKWTYYYYYMMKWS 39
QY 289 TATATGATGCTCTCGATATATTTTGATAAT 318
Db 38 KMTYSMMSYARKCWRKYAKTYTWTCMT 9

RESULT 6
US-09-949-016-13667
; Sequence 13667, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13667
; LENGTH: 89892
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(89892)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13667

Query Match 9.8%; Score 37.4; DB 3; Length 89892;
Best Local Similarity 53.0%; Pred. No. 3.6;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 224 AGACCTTGAATCTTAACCTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAAAT 283
Db 66602 AGATATGCACTCTCTATATAGACATAGATATAAAATATATGACATATTTAAATAAAT 66661
QY 284 ATCTATATATGATGCTCTCGATATATTTTGATAATCTTCTAGTATTTTGAGAAATCT 343
Db 66662 TTATATATATGATGATACATATATTAATTATACATATTAATTATATATATACATTAT 66721
QY 344 CTCAAAAAGTCTTGTGAAGTTATATTTCTTT 374
Db 66722 TTTAATAAATTATTATTATTATGAGCCTT 66752

RESULT 7
US-09-949-016-12505/c
; Sequence 12505, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12505
; LENGTH: 670690
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(670690)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12505

Query Match 9.8%; Score 37.4; DB 3; Length 670690;
Best Local Similarity 53.0%; Pred. No. 7.8;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 220 TACAAGACCTTCGAATCTTAACCTCAAGAAGCATCAATGACTTTGACACAAGGTATATACT 279
Db 544773 TAATAGACATTTTAAGCACAGCAAAATTAATTGTACCTTTAAATATGATAATGAAT 544714
QY 280 AAATATCTATATATGATGCTCTCGATATATTTTGATAATCAATCTAGTATTTTGAGAAA 339
Db 544713 GAAGCTCTATATGCCAATTTGTATGATATATATCAGCATTCCTCAATTAAGCACATTTCTTCGATA 544654
QY 340 TTCTCTCAAAAAGTCTTGTGTAAGTTATATTT 370
Db 544653 TATTCTGCAAAAAGTCTTGTGTTCTAAAGTT 544623

RESULT 8
US-09-949-016-14207/c
; Sequence 14207, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14207
; LENGTH: 670690
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(670690)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14207

Query Match 9.8%; Score 37.4; DB 3; Length 670690;
Best Local Similarity 53.0%; Pred. No. 7.8;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 220 TACAAGACCTTCGAATCTTAACCTCAAGAAGCATCAATGACTTTGACACAAGGTATATACT 279
```

Db 544773 TAATAGACATTTTAAGCACAGCAAAATTAATATTGTACCTTTAAATATGATTAATGAAC 544714
Qy 280 AATATCTATATATGATCTCTCGATATATTTTGTAAATCATCTAGTGTAGTTTGAGAAA 339
Db 544713 GAAGCTCTATATGCCAATGTATGATATATATACAGCATTCCTCCATTAAGCACATCTTCGATA 544654
Qy 340 TTCTCTCAAAAAGTCTCTGTAAGTTATATTT 370
Db 544653 TATCTGCAAAAAGTCTCTGTTCTTAAAGTT 544623

RESULT 9
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PaeI promoters
US-09-806-708B-22

Query Match 9.7%; Score 37; DB 3; Length 1141;
Best Local Similarity 14.7%; Pred. No. 0.84;
Matches 44; Conservative 91; Mismatches 164; Indels 0; Gaps 0;
Qy 71 TGGCAGTTTACGACAGGACACACCCGACCGGCAATGTGGCAAAAGCTGTCGGAG 130
Db 956 TMBGWADTAGKNNNNNNWTDVRMAKAKNNNNNNYATACYNRAATNNKMAHWM 897
Qy 131 GGAATACTGTAGAGAGTGAAGCGCACATATGACATCTCGTGAGGATCTCATCAACA 190
Db 896 KWTGHAHSKRTRHHTRCRTKYNNNNNNARTVYVYHHAARRMNAWMTRTNNNNNNN 837
Qy 191 TCGAGACTGTCGTGTCCTTCGCCAATTACAAGACCTTCGAATCTAACTCAAGAGCA 250
Db 836 NNACRNRTRTWABWKHSWNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 777
Qy 251 TCAATGACTTTGACACAGGATATATACTAAATATCTATATATGATCTCTCGATATAT 310
Db 776 TCNNYMAAATTTTDCYKWTWNTWDMWTBTTTNNNNNNNNNNNNNNNNNNNNNNNN 717
Qy 311 TTGATAATCATCTTAGTCAATTTTGAGAAATCTCTCAAAAAGTCTCTGTAAGTATAT 369
Db 716 NNNNNNNWKAYYAHATNNWGCWNNNTDARRNTNTVMRRRWMNTKTRWYSTTRRHHT 658

RESULT 10
US-08-447-500-11
; Sequence 11, Application US/08447500
; Patent No. 5627064
; GENERAL INFORMATION:
; APPLICANT: Hoechst, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA

; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 294..1385
US-08-447-500-11

Query Match 9.7%; Score 36.8; DB 2; Length 1689;
Best Local Similarity 56.7%; Pred. No. 1.1;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
Qy 246 AAGCATCAATGACTTTTGACACAGGTATATACTAAATATCTATATATGATCTCGAT 305
Db 1551 AAACCCAAATGCTTGTTCATATATATGATGATTTGTATATACATATATGTTGT 1610
Qy 306 ATATTTGATATCATCTTAGTCAATTTGAGAAATCTCTCAAAAAGTCTCTGTAAGTTA 365
Db 1611 ATATTTATATCATCTCTTGGGATTTGGGTCAATTTTTTAAACAACATGCAATCTTT 1670

RESULT 11
US-08-453-866-11
; Sequence 11, Application US/08453866
; Patent No. 5756289
; GENERAL INFORMATION:
; APPLICANT: Hoechst, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,866
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/008,001
FILING DATE: 20-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-2458
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1689 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE: Protein Kinase
CLONE: Protein Kinase
FEATURE:
NAME/KEY: CDS
LOCATION: 294..1385
US-08-453-866-11

Query Match 9.7%; Score 36.8; DB 2; Length 1689;
Best Local Similarity 56.7%; Pred. No. 1.1;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 246 AAGCATCAATGACTTTGACACAGGTATATACTAAATATCTATATATGATGCTCTCGAT 305
DB 1551 AAACCCAAATGCTTGTTCATATATATATATATATATGATGATTTGTATATACATATATGTTGT 1610
QY 306 ATATTGTAATCAATCTAGTATTTTCAGAAATTTCTCAAAAAGTTCTTGTAAAGTTA 365
DB 1611 ATATTATATCAATCTCTGGATTTGGGTCATTTTAACTCACTCTTTTAA 1670

RESULT 12
US-08-454-097-9
Sequence 9, Application US/08454097
Patent No. 5686412
GENERAL INFORMATION:
APPLICANT: Hoeckstra, Merl F.
TITLE OF INVENTION: Protein Kinases
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,097
FILING DATE: 30-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/185,359
FILING DATE: 21-JAN-1994
APPLICATION NUMBER: US 08/008,001
FILING DATE: 21-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,783
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5686412and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31853
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 297..1388
US-08-454-097-9

Query Match 9.7%; Score 36.8; DB 2; Length 2385;
Best Local Similarity 56.7%; Pred. No. 1.3;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 246 AAGCATCAATGACTTTGACACAGGTATATACTAAATATCTATATATGATGCTCTCGAT 305
DB 1554 AAACCCAAATGCTTGTTCATATATATATATATATATGATGATTTGTATATACATATATGTTGT 1613
QY 306 ATATTGTAATCAATCTAGTATTTTCAGAAATTTCTCAAAAAGTTCTTGTAAAGTTA 365
DB 1614 ATATTATATCAATCTCTGGATTTGGGTCATTTTAACTCACTCTTTTAA 1673

RESULT 13
US-08-468-036-41
Sequence 41, Application US/08468036
Patent No. 5728806
GENERAL INFORMATION:
APPLICANT: Demaggio, Anthony J.
TITLE OF INVENTION: Materials and Methods Relating to Proteins that
INTERACT WITH CASEIN KINASE I
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,036
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,605
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5728806and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31784
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-468-036-41

```
Query Match          9.7%; Score 36.8; DB 2; Length 2385;
Best Local Similarity 56.7%; Pred. No. 1.3;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 246 AAGCATCAATGACTTTTGACACAAAGGTATATAAATACTATATATCTATATATGATGCTCTCGAT 305
Db 1554 AAACCCNAATGCTCTTGCATATATATGATATATGATATTTGTTATATACATATATGTTGT 1613

QY 306 ATATTTTGATAATCATCTAGTAGTATTTGAGAAATCTCTCAAAAAGTTCTTCTGAAGTTA 365
Db 1614 ATATTTATATCAITTCCTCTTGGGATTTTGGGTCATTTTTTAACAACGTCATCTTTTAA 1673

RESULT 14
US-08-376-843-41
; Sequence 41, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; CORRESPONDENCE ADDRESS: 53
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5846764and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-376-843-41

Query Match          9.7%; Score 36.8; DB 2; Length 2385;
Best Local Similarity 56.7%; Pred. No. 1.3;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 246 AAGCATCAATGACTTTTGACACAAAGGTATATAAATACTATATATCTATATATGATGCTCTCGAT 305
Db 1554 AAACCCNAATGCTCTTGCATATATATGATATATGATATTTGTTATATACATATATGTTGT 1613

QY 306 ATATTTTGATAATCATCTAGTAGTATTTGAGAAATCTCTCAAAAAGTTCTTCTGAAGTTA 365
Db 1614 ATATTTATATCAITTCCTCTTGGGATTTTGGGTCATTTTTTAACAACGTCATCTTTTAA 1673
```

```
RESULT 15
US-08-185-359-9
; Sequence 9, Application US/08185359
; Patent No. 6060296
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Protein Kinases
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS: 57
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/185,359
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,783
; FILING DATE: 03-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6060296and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 297..1388
US-08-185-359-9

Query Match          9.7%; Score 36.8; DB 3; Length 2385;
Best Local Similarity 56.7%; Pred. No. 1.3;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 246 AAGCATCAATGACTTTTGACACAAAGGTATATAAATACTATATATCTATATATGATGCTCTCGAT 305
Db 1554 AAACCCNAATGCTCTTGCATATATATGATATATGATATTTGTTATATACATATATGTTGT 1613

QY 306 ATATTTTGATAATCATCTAGTAGTATTTGAGAAATCTCTCAAAAAGTTCTTCTGAAGTTA 365
Db 1614 ATATTTATATCAITTCCTCTTGGGATTTTGGGTCATTTTTTAACAACGTCATCTTTTAA 1673

Search completed: March 10, 2006, 18:45:53
Job time : 140 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model.

Run on: March 10, 2006, 18:32:51 ; Search time 2647 Seconds

(without alignments)
6734.370 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttctggttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

10: gb_est10:*

11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| 1 | 181 | 47.5 | 256 | 9 | CC179487 | CC179487 SALK_0699 |
| 2 | 146.8 | 38.5 | 793 | 9 | BH947619 | BH947619 obu80f11 |
| 3 | 137.2 | 36.0 | 446 | 8 | L38243 | L38243 BNAF0581E M |
| 4 | 128.8 | 33.8 | 456 | 7 | CO052301 | CO052301 Mdfw20570 |
| 5 | 124.2 | 32.6 | 632 | 6 | CF230621 | CF230621 PtaC0010H |
| 6 | 122.6 | 32.2 | 662 | 7 | CV277080 | CV277080 WS0142.B2 |
| 7 | 121.8 | 32.0 | 728 | 5 | BU890694 | BU890694 P040E03 P |
| 8 | 121.2 | 31.8 | 457 | 6 | CD669972 | CD669972 eepic.pk0 |
| 9 | 121 | 31.8 | 516 | 7 | CV278823 | CV278823 WS0147.B2 |
| 10 | 120.4 | 31.8 | 588 | 7 | CV283381 | CV283381 WS0187.B2 |
| 11 | 120.4 | 31.6 | 624 | 7 | CV232994 | CV232994 WS0199.B2 |
| 12 | 120.4 | 31.6 | 627 | 7 | CV230387 | CV230387 WS0191.B |
| 13 | 119.8 | 31.4 | 626 | 6 | CB920052 | CB920052 VSD058F11 |
| 14 | 119.8 | 31.4 | 662 | 6 | CB919185 | CB919185 VVD042D03 |
| 15 | 118.2 | 31.0 | 474 | 5 | BU668323 | BU668323 MC01026A1 |
| 16 | 117.8 | 30.9 | 349 | 5 | BU822353 | BU822353 UB36DP808 |
| 17 | 117 | 30.7 | 513 | 7 | CO997831 | CO997831 pam01-16m |
| 18 | 115.4 | 30.3 | 599 | 8 | CV882244 | CV882244 Mdlv3_403 |
| 19 | 115.4 | 30.3 | 653 | 8 | DR997716 | DR997716 Mdfb8004M |
| 20 | 115.4 | 30.3 | 659 | 8 | DR990507 | DR990507 Mdlr7005H |
| 21 | 115.4 | 30.3 | 660 | 8 | DR995864 | DR995864 Mdaeg9011J |
| 22 | 115.4 | 30.3 | 688 | 7 | CV186968 | CV186968 Mdlv2_401 |

| | | | | | | |
|------|-------|------|-----|---|----------|---------------------|
| 23 | 113.6 | 29.8 | 629 | 8 | DR994824 | DR994824 Mdaeg9007G |
| 24 | 113.4 | 29.8 | 577 | 1 | AJ768009 | AJ768009 AJ768009 |
| 25 | 112.4 | 29.5 | 460 | 2 | BG726181 | BG726181 BACCA01.0 |
| 26 | 112.2 | 29.4 | 497 | 6 | CF603713 | CF603713 BACCA01.0 |
| c 27 | 111.4 | 29.2 | 476 | 6 | CB035850 | CB035850 VVA008A06 |
| c 28 | 110.6 | 29.0 | 611 | 6 | CF373337 | CF373337 CSECS086D |
| 29 | 110 | 28.9 | 749 | 6 | CB972859 | CB972859 CAB30001 |
| 30 | 109.8 | 28.8 | 492 | 6 | CB002633 | CB002633 VAB019A06 |
| 31 | 109.8 | 28.8 | 492 | 6 | CB002772 | CB002772 VVB030F07 |
| 32 | 109.8 | 28.8 | 492 | 6 | CB004008 | CB004008 VVB034G07 |
| 33 | 109.8 | 28.8 | 509 | 6 | CB344615 | CB344615 CA48EN000 |
| 34 | 109.8 | 28.8 | 513 | 6 | CB342316 | CB342316 CA32EN000 |
| 35 | 109.8 | 28.8 | 528 | 6 | CA811699 | CA811699 CA41LN021 |
| 36 | 109.8 | 28.8 | 585 | 3 | BM437293 | BM437293 VVA017C08 |
| 37 | 109.8 | 28.8 | 613 | 6 | CB345061 | CB345061 CA48EN000 |
| 38 | 109.8 | 28.8 | 618 | 6 | CB344529 | CB344529 CA48EN000 |
| 39 | 109.8 | 28.8 | 623 | 6 | CB343483 | CB343483 CA32EN000 |
| c 40 | 109.8 | 28.8 | 636 | 6 | CB343539 | CB343539 CA32EN000 |
| 41 | 109.8 | 28.8 | 637 | 6 | CA811769 | CA811769 CA41LN031 |
| 42 | 109.8 | 28.8 | 697 | 6 | CB343846 | CB343846 CA32EN000 |
| 43 | 108.8 | 28.6 | 556 | 1 | AJ320038 | AJ320038 AJ320038 |
| c 44 | 108.8 | 28.6 | 557 | 1 | AJ320048 | AJ320048 AJ320048 |
| 45 | 108.8 | 28.6 | 558 | 1 | AJ320067 | AJ320067 AJ320067 |

ALIGNMENTS

RESULT 1
CC179487

LOCUS

DEFINITION

Arabisopsis thaliana genomic clone SALK_069941.39.90.X, genomic

survey sequence.

ACCESSION

CC179487

VERSION

GSS.

KEYWORDS

Arabidopsis thaliana (thale cress)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salg.edu

This is single pass sequence recovered from the left border of

TDNA. This sequence lies within 300 bases of the 5' end of

ATg75250.

Class: TDNA tagged.

Location/Qualifiers

1..256

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="SALK_069941.39.90.x"

/cdate="19980113"

/note="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

```

be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match          47.5%; Score 181; DB 9; Length 256;
Best Local Similarity 94.9%; Pred. No. 9.7e-38;
Matches 187; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTACCATGGACGTTTATAGTCAAAACAAGATGTTTC 60
Db 60 ATGGCGTCAAACTCAAAAGTTCAATCTACCATGGACGTTTATAGTCAAAACAAGATGTTTC 119

QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGTCGCAAAA 120
Db 120 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGTCGCAAAA 179

QY 121 GCTGTCCGAGGGAAGAACTGTAGAAGAGTGAAGCGCCACTATGACATTTCTGTCGAGGAT 180
Db 180 GCTGTCCGAGGGAAGAACTGTATAGAAGTGAAGCGCCCTCTATGACATTTCTGTCGAGAA 239

QY 181 CTCATCAACATCCAGAC 197
Db 240 CTCATCAACATCCAGAC 256

RESULT 2
BH947619/c
LOCUS          793 bp DNA linear GSS 01-OCT-2002
DEFINITION    obu80f11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
               sequence.
ACCESSION     BH947619
VERSION       BH947619.1 GI:23427679
KEYWORDS      GSS.
SOURCE        Brassica oleracea
ORGANISM      Brassica oleracea
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 793)
AUTHORS       Delehaanty,K., Fellw,G., Fulton,L., McCombie,W.R., Miner,T.,
               Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE         Whole genome shotgun reads from Brassica oleracea
JOURNAL       Unpublished (2002)
COMMENT       Contact: Richard K. Wilson
               Genome Sequencing Center
               Washington University School of Medicine
               Email: submissions@watson.wustl.edu
               Plate: obu80 row: f column: 11
               Seq primer: -21UPpOT forward
               Class: shotgun
               High quality sequence start: 112
               High quality sequence stop: 551.
               Location/Qualifiers
                   1..793
                       /organism="Brassica oleracea"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:3712"
                       /clone_lib="B.oleracea002"
                       /notes="Vector: pOTw13; Whole genome shotgun library from
               flowering buds. DNA was purified from a crude nuclear
               prep using Brassica oleracea TO1000PH3 buds provided by
               Thomas Osborn at the University of Wisconsin. Genomic
               DNA was provided by Pablo Rabinowicz (CSHL) and the
               shotgun library prepared at Washington University Genome
               Sequencing Center."

ORIGIN
Query Match          38.5%; Score 146.8; DB 9; Length 793;
Best Local Similarity 90.2%; Pred. No. 1.8e-28;
Matches 157; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 107 ACAATGGCAAAAGCTGTCGGAGGGAAGAACTGTAGAAGAGTGAAGCGCCACTATGACA 166
Db 545 AAATCTCTCAACGCTGTCGGAGGGAAGAACTGTCAGACGAGGCGATCACTATGACA 486

```

```

QY 167 TTCTCGTCGAGGATCTCATCAACATCGACACTGGTGGTCCCTTTGCCCAATTACAAGA 226
Db 485 TTCTCGTCGAGGATCTCATCAACATCGAAACTGGTGGTCCCTTTGCCCAATTACAAGA 426

QY 227 CCTTCGAATCTAACTCAAGAAGCATCAATGACTTTTGACACAAGGTATATAACTA 280
Db 425 CTTTCGATTTCTAACTCAAGAGGATCAATGACTTCGACACGAGGTATATAGCTA 372

RESULT 3
L38243
LOCUS          446 bp mRNA linear EST 03-JUL-1995
DEFINITION    ENAF0581E Mustard flower buds Brassica rapa cDNA, mRNA sequence.
ACCESSION     L38243
VERSION       L38243.1 GI:887283
KEYWORDS      EST.
SOURCE        Brassica rapa (Brassica campestris)
ORGANISM      Brassica rapa
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 446)
AUTHORS       Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H.,
               Hwang,I. and Cho,M.J.
TITLE         Expressed sequence tags of Chinese cabbage flower bud cDNA
JOURNAL       Plant Physiol. 111 (2), 577-588 (1996)
PUBMED        8787028
COMMENT       Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.
               Plant Molecular Biology and Biotechnology Research Center
               Email: pmbr@cnongae.gnu.ac.kr.
               Location/Qualifiers
                   1..446
                       /organism="Brassica rapa"
                       /mol_type="mRNA"
                       /strain="pekinensis"
                       /db_xref="taxon:3711"
                       /clone_lib="Mustard flower buds"
                       /note="Devel_stage = flower bud "

ORIGIN
Query Match          36.0%; Score 137.2; DB 8; Length 446;
Best Local Similarity 71.4%; Pred. No. 6.2e-26;
Matches 197; Conservative 0; Mismatches 73; Indels 6; Gaps 1;

QY 12 CTCAGAAGTTCAATCTCACCATGGACGTTTATAGTCAAAACAAGATTTTCGAGAGGCGCTT 71
Db 81 CTCAAGTTCTATGAGCTCTCTTGGACGCTTAAGCAAAATAAGATATTCGAAAGGCGCTT 140

QY 72 GGCAGTTTACGACNAGGACACACCCGACGATGGCACATTCGTTCGAGAGTCTCATCAACAT 131
Db 141 AGCCGTTTACGACAAAGACACTCCAGACGTTTGGCAAAACGTCGGCAAGCCGTCGGAA 200

QY 132 GAAACTGTAGAAGAGTGAAGCGCCACTATGACATTCGTTCGAGAGTCTCATCAACAT 191
Db 201 TAAATCGGACAGGAAGTCAAAGCTCAGTACGATATTCGTTCGAGAGTCTCATGACAT 260

QY 192 CGAGACTGGTCGTGTCCTTTGCCCAATTACAAGACCTTCGA-----ATCTAACTCAAG 245
Db 261 CGAGCAAGACTTAGTCCCTTTGGCTAAATACAAGACCGTCGATGTTGGAAATAAATCAAG 320

QY 246 AAGCATCATGACTTTTGACACAGGTATATAACTAA 281
Db 321 AGGCATCAACGGTTACGGTTTGAGGTTTAATGAAAA 356

RESULT 4
CO052301
LOCUS          456 bp mRNA linear EST 14-JUN-2004
DEFINITION    Mdfw2057o18.y1 Mdfw Malus x domestica cDNA clone Mdfw2057o18.5;
               similar to TR:Q9S1J5 Q9S1J5 AT2G21650 PROTEIN. 1, mRNA sequence.
ACCESSION     CO052301
VERSION       CO052301.1 GI:48693766

```

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST.
 Malus x domestica
 Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
 1 (bases 1 to 456)

Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A.,
 Aldwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Orvis, K.,
 Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T.,
 Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
 Tsagareishvili, R., Kennedy, S., Waterston, K., and Wilson, R.
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)

Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu

Library materials provided by: Schuyler S. Korban Library
 constructed by: A. Hernandez / K. Gasic Library sequenced by:
 Washington University Genome Sequencing Center

WashU EST name: aah28h09.y1

Seq primer: -40UP from Gibco

High quality sequence stop: 456.

Location/Qualifiers

FEATURES

source

1..456
 /organism="Malus x domestica"
 /mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/clone="WdFw2057018"

/tissue_type="Flower"

/lab_host="DH10B ampicillin resistant"

/clone_lib="WdFw"

/note="Vector: pBluescript II SK (+); Site 1: NotI;
 Site 2: EcoRI; Total RNA was extracted separately from
 each floral stage (bud, balloon, open and after
 pollination), using the 'pine tree' method. Poly(A)+mRNA
 was isolated twice from total RNA from each stage using
 the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse
 transcribed into double stranded cDNA using a modified
 oligo18(dT) primer with an identifying tag sequence (see
 table below). cDNA's from different stages were pooled in
 equal amounts before adaptor ligation. Tag identification
 when sequencing from 5' end: Stage 1 (bud) insert
 18(A)TCGGA; Stage 2 (balloon) insert 18(A)TGCGA; Stage 3
 (open) insert 18(A)TGCGT; Stage 4 (afterpollination)
 insert 18(A)TCGCT. Tag identification when sequencing from
 3' end: Stage 1 (bud) TCGGA18(T) insert; Stage 2 (balloon)
 TCGGA18(T) insert; Stage 3 (open) ACGCA18(T) insert; Stage
 4 (afterpollination) ACCGA18(T) insert. Double stranded
 cDNAs were size selected (more than 450 bp), adaptor
 with EcoRI adaptors at both ends and then digested with
 NotI. The cDNAs were then directionally cloned into
 EcoRI-NotI digested pBS II SK(+) phagemid
 vector(Stratagene). Identification of adaptors and tags in
 5'-end sequenced clones: (Vector) ... TAAAGCTT(End
 Vector) (Start EcoRI adaptor)GATATCGAATTCATGTGTGGG (End
 EcoRI adaptor) (Start Insert) ... AAAAAAAAAAAAAAAAAA(End
 Insert) (Start Tag)TCGGA(End Tag) (Start
 NotI site)/vector/GCGCGCCACCGGG... The total number of
 white colony forming units (cfu) in the primary library
 before amplification was 1.1x10⁶ cfu (colony forming
 units). The background of empty clones was less than 1%.
 Inserts ranged from 0.5kb to 3 kb, as determined by PCR.
 Purified plasmid DNA from the primary library was
 converted to single-stranded circles and used as a template
 for PCR amplification using the T7 and T3 priming sites
 flanking the cloned cDNA inserts. The purified PCR
 products, representing the entire cloned cDNA population,

were used as a driver for normalization. Hybridization
 between the single-stranded library and the PCR products
 was carried out for 44 hours at 30C. Unhybridized
 single-stranded DNA circles were separated from hybridized
 DNA rendered partially double-stranded and electroporated
 into DH10B cells to generate the normalized library. The
 total number of clones with insert was 9x10⁶ cfu.
 Background of empty clones was less than 1%."

ORIGIN

Query Match 33.8%; Score 128.8; DB 7; Length 456;
 Best Local Similarity 69.4%; Pred. No. 1.1e-23;
 Matches 175; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 3 | GGCGTCAAACTCAAGAGTTCAATCTCACCATTGGAGCTTTAGTCAAAAAGATGTTTCA | 62 |
| DB | 56 | GGTTTCTTCTCGAGGAATTTCAACTCTCTGGACGCGGAGGAGAACAAAGCTGTTTCA | 115 |
| QY | 63 | GAGGCGCTTGGCAGTTTACGACAAAGGACACACCCGACCGATGGCAATGTGGCAAAAGC | 122 |
| DB | 116 | GAAGGCGCTGGCGGTAGCACAAGGACAGCAGGACCGCTGGCATATGTGGCAAGGC | 175 |
| QY | 123 | TGTCGAGGAGAAAACCTGTAGAAAGTGAAGCGCCATATGACATTTCTCGTCCAGGATCT | 182 |
| DB | 176 | CGTCGCGGAAAATCTGCGGAGGAAGTAAAGAGGCACTACGAGATTTCTGGTGCATGATCT | 235 |
| QY | 183 | CATCAACATCGAGACTGGTCTGTCCTTTGGCCCAATTACAGACCTTGAATCTTAACTC | 242 |
| DB | 236 | CATGCACATAGAAATCCGCCCATGTCCCTATATCAAAATATTAAAGTCGCGCGGACGCGG | 295 |
| QY | 243 | AAGAGCATCAA | 254 |
| DB | 296 | CAGCGGACGCAA | 307 |

RESULT 5

CF230621

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CF230621 632 bp mRNA linear EST 05-AUG-2003
 PtAC0010H11H115 Poplar cDNA library from cambial zone Populus alba
 x Populus tremula cDNA 5', mRNA sequence.

CF230621
 CF230621.1 GI:33450050
 EST.

Populus alba x Populus tremula
 Populus alba x Populus tremula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.

1 (bases 1 to 632)
 Dejardin, A., Leple, J.-C., Lesage-Descauses, M.-C., Costa, G. and
 Pilate, G.

Expressed sequence tags from poplar wood tissues - A comparative
 analysis from multiple libraries
 Plant Biol. 6 (1), 55-64 (2004)

15095135

Contact: Leple JC

Unit of Forest Improvement, Genetics and Physiology

National Institute for Agricultural Research (INRA)

Domaine de Limere, BP20619 ARDON, 45166 OLIVET CEDEX, FRANCE

Tel: 33 02 38 41 78 00

Fax: 33 02 38 41 78 79

Email: Jean-Charles.Leple@leons.inra.fr

PCR Primers

FORWARD: TriplexA 5' CTCGGAGAGCGCGCATTTGTG 3'

BACKWARD: TriplexB1 5' ATACGACTCTATAGGCGCA 3'

Plate: PAc0010 row: H column: 11

Seq primer: TriplexA 5' CTCGGAGAGCGCGCATTTGTG 3'.

Location/Qualifiers

1..632

/organism="Populus alba x Populus tremula"

/mol_type="mRNA"

/strain="clone INRA 717-1-B4"

/db_xref="taxon:80863"

FEATURES

source

```

/sex="female"
/tissue_type="cambial zone harvested on the bark side"
/dev_stage="3-years-old poplar trees grown in the nursery"
/clone_lib="Poplar cDNA library from cambial zone"
/notes="A composite cDNA library was made with mRNA
isolated from opposite and tension wood tissues
corresponding to the cambial zone collected on the bark
side after debarking the stem. In this respect, in
addition to cambium cDNA, this library also contains very
young phloem and very young xylem cDNA. The sampling was
done on 3 different tilted trees grown in the nursery.
cDNA were cloned in an oriented way into SfiI (A and B)
restriction sites. A one-step conversion of lambda
triplex2 to the corresponding triplex2 plasmid was done
via site-specific recombination at loxp sites (Clontech ;
SMART cDNA library construction kit). cDNA inserts were
PCR amplified using flanking primers and then sequenced on
a ABI3100 Genetic Analyser (Applied Biosystem)"

```

ORIGIN

```

Query Match 32.6%; Score 124.2; DB 6; Length 632;
Best Local Similarity 74.6%; Pred. No. 2e-22;
Matches 156; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACCATTGAGCTTTAGTCAAAAACAAGATTTCGAGAGGGCCTTGGCAGTT 78
    |||||
Db 211 AATTCTCTCACTCTTGGACACCTTAAGCAAAACAAGATCTTCGAAAAGGCCCTGGCTTTA 270
    |||||

QY 79 TAGCACAAGGACACACCCGACCGATGCGACAATGTGCAAAAGCTGTGCGAGGGGAAAAC 138
    |||||
Db 271 TAGCACAAGGACACCCCTGACCCCTGGCATTAATGTTGCCACACCTGTGGGTGGGAATCT 330

QY 139 GTAGAAGAAGTGAAGCGCCACTATGACATTTCTCGTCCGAGGATCTCATCAACATCGAGACT 198
    |||||
Db 331 GCAGACGAGTGAAGAGGCACTATGAGATTCTCATCAAGGATGTCAGGGAATCGAGTCT 390

QY 199 GGTGCGTGTCCTTGGCCCAATTACAAGAC 227
    |||||
Db 391 GGCAGAGTTCCATTCCCTTAATTACAGTC 419

```

RESULT 6

```

CV277080/c
LOCUS CV277080 662 bp mRNA linear EST 22-SEP-2004
DEFINITION WS0142.B21_J13 PTX-D-IL-A-5 Populus trichocarpa x Populus deltoides
cDNA clone WS0142_J13 3', mRNA sequence.

```

ACCESSION

```

VERSION CV277080.1 GI:52530055
KEYWORDS EST.
SOURCE Populus trichocarpa x Populus deltoides
ORGANISM Populus trichocarpa x Populus deltoides

```

```

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.

```

REFERENCE

```

AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS0142 row: J column: 13

```

TITLE

```

JOURNAL The poplar transcriptome: Analysis of expressed sequence tags from
COMMENT multiple cDNA libraries
Unpublished (2004)

```

```

High quality sequence stop: 662
POLYA=yes.

```

FEATURES

source

```

Location/Qualifiers
1..662
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"
/clone="WS0142_J13"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-IL-A-5"
/notes="Vector: pBluescript II SK (+) XR; Site 1: ECORI (5'
end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees
two metres in height and grown under greenhouse conditions
were exposed to continuous feeding by Malacosoma disstria
Hubner (forest tent caterpillar) mid-instar larvae caged
on the sapling using mesh bags. Mature leaves from within
the caged region were collected 2 hours, 12 hours, 24
hours and 48 hours after the onset of treatment. mRNA was
isolated from each tissue source independently and equal
quantities of mRNA from each tissue were then pooled. cDNA
was prepared from 5 micrograms of mRNA and directionally
ligated into the pBluescript II SK (+) XR vector using the
pBluescript II XR cDNA Library Construction kit according
to manufacturer's instructions with modifications
(Stratagene). Plasmid DNA was then transformed by
electroporation into DH10B cells (Invitrogen) for
propagation."

```

ORIGIN

```

Query Match 32.2%; Score 122.6; DB 7; Length 662;
Best Local Similarity 74.2%; Pred. No. 5.5e-22;
Matches 155; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACCATTGAGCTTTAGTCAAAAACAAGATTTCGAGAGGGCCTTGGCAGTT 78
    |||||
Db 481 AATTCTCTCACTCTTGGACACCTTAAGCAAAACAAGTTCGAAAAGGCCCTGGCTTTA 422

QY 79 TAGCACAAGGACACACCCGACCGATGCGACAATGTGGCAAAAGCTGTGCGAGGGGAAAAC 138
    |||||
Db 421 TATGACAGGACACCCCTGACCGCTGGCATGATGTTGCCAAAGCTGTGGGTGGGAATCT 362

QY 139 GTAGAAGAAGTGAAGCGCCACTATGACATTTCTCGTCCGAGGATCTCATCAACATCGAGACT 198
    |||||
Db 361 GCAGAGGAAGTGAAGAGGCACTATGAGATTCTCATCAAGGATGTCAGGGAATCGAGTCT 302

QY 199 GGTGCGTGTCCTTGGCCCAATTACAAGAC 227
    |||||
Db 301 GGCAGAGTTCCATTCCCTTAATTACAGTC 273

```

RESULT 7

```

LOCUS BU890694 728 bp mRNA linear EST 17-OCT-2002
DEFINITION P049E03 Populus petioles cDNA library Populus tremula cDNA 5 prime,
mRNA sequence.

```

ACCESSION

```

VERSION BU890694
BU890694.1 GI:24101759
KEYWORDS EST.
SOURCE Populus tremula
ORGANISM Populus tremula

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 728)
Unneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
Unpublished (2002)
Contact: BHALERAO RUPALI R.
Umea Plant Science Center
Department of Plant Physiology

```

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation.

ORIGIN

Query Match 31.8%; Score 121; DB 7; Length 516;
 Best Local Similarity 73.7%; Pred. No. 1.4e-21;
 Matches 154; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACATGGAGCTTTAGTCAAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTT 78
 |||||
 Db 481 AATTCTCTCACTCTTGGACACCTAAGCAAAACAACATATTGAAAAAGGCCCTGGCTTTA 422
 |||||

QY 79 TAGCAAGACGACACCCGACCGATGCGCAATGTGCAAAAGCTGTCGAGGGGAAACT 138
 |||||
 Db 421 CATGCAAGACACCCCTGACCGCTGGCATAATGTTGCCAAAGCTGGGTGGGAATCT 362
 |||||

QY 139 GTAGAAGAGTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACT 198
 |||||
 Db 361 GCAGAGGAGTGAAGAGGCATATGAGATTTCTCATCAAGGATGTCAGGGAATTGAGTCT 302
 |||||

QY 199 GTCGTGTCCTTTGCCAATTACAAGAC 227
 |||||
 Db 301 GGCAGAGTTCCATTCCTCAATTACAGGTC 273
 |||||

RESULT 10

CV283381/c
 LOCUS
 DEFINITION WS0187.B21.C24 PTx-D-IL-N-A-9 Populus trichocarpa x Populus deltoides cDNA clone WS0187_C24 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CV283381.1 GI:52536356

Populus trichocarpa x Populus deltoides
 Populus trichocarpa x Populus deltoides
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 588)

Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y., Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G., Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A., Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A., Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and Bohlmann, J.

TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

JOURNAL Unpublished (2004)

COMMENT Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS0187 row: C column: 24
 High quality sequence stop: 588
 POLYA=Yes.

FEATURES
 source Location/Qualifiers
 1..588 /organism="Populus trichocarpa x Populus deltoides"

/mol_type="mRNA"
 /cultivar="Hil-11"
 /db_xref="taxon:3695"
 /clone="WS0187_C24"
 /sex="Male"
 /lab_host="E. coli DH10B T1 phage resistant cells"
 /clone_lib="PTXD-IL-N-A-9"
 /note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma disstria Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods (Bonaldo M.F. et al. (1996) Genome Research 6(9):791) in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 31.8%; Score 121; DB 7; Length 588;
 Best Local Similarity 73.7%; Pred. No. 1.5e-21;
 Matches 154; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACATGGAGCTTTAGTCAAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTT 78
 |||||
 Db 432 AATTCTCTCACTCTTGGACACCTAAGCAAAACAACATATTGAAAAAGGCCCTGGCTTTA 373
 |||||

QY 79 TAGCAAGACGACACCCGACCGATGCGCAATGTGCAAAAGCTGTCGAGGGGAAACT 138
 |||||
 Db 372 CATGCAAGACACCCCTGACCGCTGGCATAATGTTGCCAAAGCTGGGTGGGAATCT 313
 |||||

QY 139 GTAGAAGAGTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACT 198
 |||||
 Db 312 GCAGAGGAGTGAAGAGGCATATGAGATTTCTCATCAAGGATGTCAGGGAATTGAGTCT 253
 |||||

QY 199 GTCGTGTCCTTTGCCAATTACAAGAC 227
 |||||
 Db 252 GGCAGAGTTCCATTCCTCAATTACAGGTC 224
 |||||

RESULT 11

CV232994/c
 LOCUS
 DEFINITION WS0199.B21.E24 PT-DX-N-A-10 Populus trichocarpa cDNA clone WS0199_E24 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CV232994.1 GI:52389099

Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
 Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 624)

Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y., Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G., Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A., Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A., Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and Bohlmann, J.

TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

JOURNAL Unpublished (2004)

| | Query Match | 31.6% | Score 120.4; | DB 7; | Length 627; |
|----|-----------------------|--|--------------------|-----------|-------------|
| | Best Local Similarity | 67.6%; | Pred. No. 2.1e-21; | | |
| | Matches 169; | Conservative 0; | Mismatches 81; | Indels 0; | Gaps 0 |
| Qy | 8 | CAAACTCAAGAAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTCGAGAGGG | 67 | | |
| Db | 569 | CATCTTCAAAATGGACCGGCTCTCTTTGGACAGCCAAACAATAAGCTATTTCGAGAAGG | 510 | | |
| Qy | 68 | CTTTGGCAGTTTACGACAAGSAGACACCCGACCGATGCGCAATGTGCGCAAAAGCTGTCG | 127 | | |
| Db | 509 | CCCTGGCTGTATACGCAAAAGACACCCAGACCGCTGSCAAATGTGGCCNAGGCCGTGG | 450 | | |
| Qy | 128 | GAGGGAAAACGTAGAAAGAGTGAAAGGCCACATATGACATTTCTCGTCGAGGATCTCATCA | 187 | | |
| Db | 449 | GTGGCAAGTCTCTGAAGAAAGTTTAAGAGGCACATATGATCGTCTCTGTGGAGGATCTCGTGT | 390 | | |
| Qy | 188 | ACATCGAGACTGGTGTGTCCTCTTGGCCCAATTACAAGACCTTTCGAATCTTAACCTCAAGAA | 247 | | |

Db 389 ACATAGATCCGCCAAGCCCTCTCGCGAATTACAAGCCCTTCGGAGCAATGGTAGAG 330
Qy 248 GCATCAATGA 257
Db 329 GACTTGTGA 320

RESULT 13
CB920052 626 bp mRNA linear EST 25-APR-2003
LOCUS VVD058F11.1 350737 An expressed sequence tag database for abiotic
DEFINITION stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
CDNA clone VVD058F11.5, mRNA sequence.

ACCESSION CB920052
VERSION CB920052.1 GI:30134714
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 626)
AUTHORS Cushman,J.C.
TITLE An expressed sequence tag database for abiotic stressed berries of
Vitis vinifera var. Chardonnay

JOURNAL
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 058 row: F column: 11
Seq primer: T3 20mer
High quality sequence stop: 626.
FEATURES
source Location/Qualifiers
1..626
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD058F11"
/tissue_type="berries"
/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Query Match 31.4%; Score 119.8; DB 6; Length 626;
Best Local Similarity 70.5%; Pred. No. 3.1e-21;
Matches 160; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 12 CTCAGAAGTTCAATCTCACCATGGAGCGTTTGTAGTCAAAACAAGATGTCGAGAGGGCGCTT 71
Db 168 CTCGCCGAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAACAAGCCCTTTGAAGAGGGCGCT 227
Qy 72 GGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAAGCTGTCCGAGG 131
Db 228 TGCAGTGTATGACAAGGACACCCCTGACCGCTGTGTACAAATGTTGCCAGGGCGCTTGTGGTGG 287
Qy 132 GAAACCTGTAGAAAGTGAAGCCCACTATGACATTCCTCGTCAGGATCTCATCAACAT 191
Db 288 GAAACCCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTTGTGTGAGGATATCAAGAGCAT 347
Qy 192 CGAGACTGGTCTGCTGCCCTTTGCCCAATTACAAGACCTTCGAATCTA 238
Db 348 CGACTCAGACAAAGTGGCCCTTCCCAATTACAAGACTACCGGAGCTA 394

RESULT 15
BU668323 474 bp mRNA linear EST 31-DEC-2003
LOCUS MC01026A10 MC01 Sesamum indicum cDNA, mRNA sequence.
DEFINITION
ACCESSION BU668323

RESULT 14
CB919185 662 bp mRNA linear EST 25-APR-2003
LOCUS VVD042D03.3 349003 An expressed sequence tag database for abiotic
DEFINITION stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
CDNA clone VVD042D03.5, mRNA sequence.

ACCESSION CB919185
VERSION CB919185.1 GI:30133846
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Vitaceae; Vitis.

REFERENCE 1 (bases 1 to 662)
AUTHORS Cushman,J.C.
TITLE An expressed sequence tag database for abiotic stressed berries of
Vitis vinifera var. Chardonnay

JOURNAL
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 042 row: D column: 03
Seq primer: T3 20mer
High quality sequence stop: 662.
FEATURES
source Location/Qualifiers
1..662
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD042D03"
/tissue_type="berries"
/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Query Match 31.4%; Score 119.8; DB 6; Length 662;
Best Local Similarity 70.5%; Pred. No. 3.1e-21;
Matches 160; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 12 CTCAGAAGTTCAATCTCACCATGGAGCGTTTGTAGTCAAAACAAGATGTCGAGAGGGCGCTT 71
Db 168 CTCGCCGAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAACAAGCCCTTTGAAGAGGGCGCT 227
Qy 72 GGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAAGCTGTCCGAGG 131
Db 228 TGCAGTGTATGACAAGGACACCCCTGACCGCTGTGTACAAATGTTGCCAGGGCGCTTGTGGTGG 287
Qy 132 GAAACCTGTAGAAAGTGAAGCGCCACTATGACATTCCTCGTCAGGATCTCATCAACAT 191
Db 288 GAAACCCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTTGTGTGAGGATATCAAGAGCAT 347
Qy 192 CGAGACTGGTCTGCTGCCCTTTGCCCAATTACAAGACCTTCGAATCTA 238
Db 348 CGACTCAGACAAAGTGGCCCTTCCCAATTACAAGACTACCGGAGCTA 394

RESULT 15
BU668323 474 bp mRNA linear EST 31-DEC-2003
LOCUS MC01026A10 MC01 Sesamum indicum cDNA, mRNA sequence.
DEFINITION
ACCESSION BU668323

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:28:27 ; Search time 400 Seconds

(without alignments)
6348.120 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattctttgtggttaa 381

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*
14: Geneseqn2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 381 | 100.0 | 381 | 10 ADF43565 | Adf43565 Thale cre |
| 2 | 273.4 | 71.8 | 294 | 14 AEA27134 | Aea27134 Stress to |
| 3 | 149.4 | 39.2 | 279 | 12 ADO62042 | Ado62042 Transcrip |
| 4 | 138.6 | 36.4 | 478 | 3 AAC34192 | Aac34192 Arabidops |
| 5 | 122.8 | 32.2 | 482 | 14 ADW16587 | Adw16587 Eucalyptu |
| 6 | 120.6 | 31.7 | 760 | 13 ADR59489 | Adr59489 Cotton cd |
| 7 | 119 | 31.2 | 543 | 13 ACN61386 | Acn61386 Cotton gy |
| 8 | 110.8 | 29.1 | 779 | 13 ADR59490 | Adr59490 Cotton cd |
| 9 | 110.8 | 28.9 | 466 | 12 ADJ43786 | Adj43786 Plant ful |
| 10 | 108.6 | 28.5 | 513 | 11 ACL30919 | Ac130919 Rice abio |
| 11 | 108.6 | 28.5 | 508 | 3 AAC35256 | Aac35256 Arabidops |
| 12 | 108.4 | 28.5 | 577 | 13 ACN59055 | Acn59055 Cotton gy |
| 13 | 107.4 | 28.2 | 505 | 13 ACN61224 | Acn61224 Cotton gy |
| 14 | 107.4 | 28.2 | 532 | 13 ACN59701 | Acn59701 Cotton gy |
| 15 | 107.4 | 28.2 | 532 | 13 ACN56695 | Acn56695 Cotton gy |
| 16 | 107.4 | 28.2 | 538 | 13 ACN56695 | Acn56695 Cotton gy |
| 17 | 107.4 | 28.2 | 542 | 13 ACN56782 | Acn56782 Cotton gy |
| 18 | 107.4 | 28.2 | 604 | 13 ADR64219 | Adr64219 Cotton cd |
| 19 | 104.8 | 27.5 | 306 | 3 AAC42715 | Aac42715 Arabidops |

| | | | | | | |
|---|----|-------|------|------|-------------|--------------------|
| c | 20 | 104.8 | 27.5 | 556 | 10 ABX57121 | Abx57121 Arabidops |
| | 21 | 104.8 | 27.5 | 578 | 10 ADD30327 | Add30327 Plant vie |
| | 22 | 104.8 | 27.5 | 578 | 12 ADI44256 | Adi44256 Plant tra |
| | 23 | 104.8 | 27.5 | 578 | 14 AEA27034 | Aea27034 Stress to |
| | 24 | 103.8 | 27.2 | 636 | 13 ADR65153 | Adr65153 Cotton cd |
| | 25 | 103.4 | 27.1 | 584 | 13 ACN54517 | Acn54517 Cotton an |
| | 26 | 99 | 26.0 | 504 | 3 AAC49819 | Aac49819 Arabidops |
| | 27 | 98.6 | 25.9 | 507 | 13 ADX14446 | Adx14446 Plant ful |
| | 28 | 92.4 | 24.3 | 513 | 10 ADB31830 | Adb31830 DNA encod |
| | 29 | 92.4 | 24.3 | 513 | 12 ADO02248 | Ado02248 Thalecres |
| | 30 | 92.4 | 24.3 | 513 | 14 AEA27102 | Aea27102 Stress to |
| | 31 | 92 | 24.1 | 294 | 3 AAC43000 | Aac43000 Arabidops |
| | 32 | 92 | 24.1 | 294 | 14 AEA27038 | Aea27038 Stress to |
| | 33 | 92 | 24.1 | 303 | 3 AAC46353 | Aac46353 Arabidops |
| | 34 | 91 | 23.9 | 2192 | 3 AAA78533 | Aaa78533 Plant SDF |
| | 35 | 87 | 22.8 | 980 | 13 ADX51987 | Adx51987 Plant ful |
| | 36 | 84.6 | 22.2 | 776 | 12 ADJ39433 | Adj39433 Plant cdn |
| | 37 | 83.6 | 21.9 | 536 | 13 ACN58748 | Acn58748 Cotton gy |
| | 38 | 81.2 | 21.3 | 563 | 13 ACN58701 | Acn58701 Cotton gy |
| | 39 | 79.6 | 20.9 | 867 | 6 ABK65275 | Abk65275 Arabidops |
| | 40 | 79.6 | 20.9 | 867 | 10 ADD30621 | Add30621 Plant yie |
| | 41 | 79.6 | 20.9 | 867 | 10 ADE31552 | Ad311552 Plant yie |
| | 42 | 79.6 | 20.9 | 867 | 12 ADI44216 | Adi44216 Plant tra |
| | 43 | 79.4 | 20.8 | 1431 | 14 ADW16598 | Adw16598 Eucalyptu |
| | 44 | 77.6 | 20.4 | 1230 | 13 ADX13094 | Adx13094 Plant ful |
| | 45 | 75 | 19.7 | 633 | 13 ADX12277 | Adx12277 Plant ful |

ALIGNMENTS

RESULT 1
ADP43565
ID ADF43565 standard; DNA; 381 BP.
XX
AC ADF43565;
XX
DT 12-FEB-2004 (first entry)
XX
DE Thale cress peronospora parasitica resistant protein coding sequence.
XX
KW transgenic plant; peronospora parasitica resistant; PPR2;
KW pathogen resistant phenotype; modified pathogen resistance; thale cress;
KW gene; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
CDS 1..381
FT /*tag= a
FT /product= "Thale cress peronospora parasitica resistant
FT protein"
XX
PN WO2003091412-A2.
XX
PD 06-NOV-2003.
XX
PF 24-APR-2003; 2003WO-US012981.
XX
PR 24-APR-2002; 2002US-0375333P.
XX
PA (AGRI-) AGRINOMICS LLC.
XX
PI Federspiel N, Lammers A, Liu XL, Bates SR, Westerlund C;
PI Fitch JR;
XX
DR WPI; 2003-865582/80.
XX
P-PSDB; ADF43566.
XX
XX New transgenic plants with increased resistance to pathogens due to
PT altered expression of Peronospora Parasitica Resistant gene (PPR2),
PT useful for generating plants with a pathogen resistance phenotype.
XX

```
PS Example 4; SEQ ID NO 1; 36pp; English.
XX
CC The invention relates to a transgenic plant that possesses a plant
CC transformation vector comprising a nucleotide sequence that encodes a
CC peronospora parasitica resistant (PPR2) protein, or a PPR2 orthologue.
CC The transgenic plant is useful in generating plants with a pathogen
CC resistance phenotype. The PPR2 nucleic acids and proteins of the
CC invention are useful in the generation of genetically modified plants
CC having a modified pathogen resistance phenotype. The present DNA sequence
CC encodes an Arabidopsis thaliana PPR2 protein of the invention.
XX
SQ Sequence 381 BP; 122 A; 76 C; 76 G; 107 T; 0 U; 0 Other;

Query Match 100.0%; Score 381; DB 10; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.2e-98;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
DB 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
DB 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
QY 121 GCTGTGCGAGGGAAAACCTGTAGAAGAGTGAAGGCCACCTATGACATTCCTCGTCGAGGAT 180
DB 121 GCTGTGCGAGGGAAAACCTGTAGAAGAGTGAAGGCCACCTATGACATTCCTCGTCGAGGAT 180
QY 181 CTGATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
DB 181 CTGATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
QY 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 300
DB 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 300
QY 301 TCGATATATTTGATTAATCAATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
DB 301 TCGATATATTTGATTAATCAATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
QY 361 AGTTATATTTCTTTGGTTTAA 381
DB 361 AGTTATATTTCTTTGGTTTAA 381

RESULT 2
AEA27134
ID AEA27134 standard; DNA; 294 BP.
XX
AC AEA27134;
XX
DT 28-JUL-2005 (first entry)
XX
DE Stress tolerant plant-related transcription factor gene SeqID975.
XX
KW transcription factor; transgenic plant; agriculture; drought resistance;
KW stress tolerance; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2005047516-A2.
XX
PD 26-MAY-2005.
XX
PF 12-NOV-2004; 2004WO-US037584.
XX
PR 13-NOV-2003; 2003US-00714887.
PR 05-DEC-2003; 2003US-0527658P.
PR 05-FEB-2004; 2004US-0542928P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX

PI Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kunimoto RW, Gutterson NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Reddie JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JM;
XX
DR WPI: 2005-372386/38.
DR P-PSDB; AEA27135.
XX
PT New transgenic plants for producing commercially or agriculturally useful
PT plants having improved tolerance to drought, shade and low nitrogen
PT conditions.
XX
PS Disclosure; SEQ ID NO 975; 407pp; English.
XX
CC This invention relates to a novel plant transcription factor
CC polypeptides, the DNA sequences which encode them and their use in
CC creating transgenic plants. The transgenic plant and methods are useful
CC for producing commercially or agriculturally useful plants having
CC improved tolerance to drought, shade and low nitrogen conditions when
CC compared to wild-type reference plants. The present sequence is that of a
CC plant transcription factor gene which was used during the development of
CC the transgenic plants of the invention.
XX
SQ Sequence 294 BP; 99 A; 65 C; 67 G; 63 T; 0 U; 0 Other;

Query Match 71.8%; Score 273.4; DB 14; Length 294;
Best Local Similarity 97.9%; Pred. No. 6.2e-68;
Matches 277; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
DB 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
DB 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
QY 121 GCTGTGCGAGGGAAAACCTGTAGAAGAGTGAAGGCCACCTATGACATTCCTCGTCGAGGAT 180
DB 121 GCTGTGCGAGGGAAAACCTGTAGAAGAGTGAAGGCCACCTATGACATTCCTCGTCGAGGAT 180
QY 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
DB 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
QY 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 283
DB 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 283

RESULT 3
ADO62042
ID ADO62042 standard; DNA; 279 BP.
XX
AC ADO62042;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2723 coding sequence, SEQ ID 509.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
```

18-SEP-2002; 2002US-0411837P.
 17-DEC-2002; 2002US-0434166P.
 24-APR-2003; 2003US-0465809P.
 (MEND-) MENDEL BIOTECHNOLOGY INC.
 Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 WPI; 2004-330163/30.
 P-PSDB; ADO62043.
 New recombinant polynucleotide encoding transcription factor
 polypeptides, useful for producing transgenic plants with advantageous
 properties compared to a reference plant.
 Claim 1; SEQ ID NO 509; 510pp; English.
 The present invention relates to novel plant transcription factor
 proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
 sequences can be used to produce transgenic plants, which overexpress
 (II), where the transgenic plant has an altered trait as compared to a
 non-transgenic plant or wild-type plant. The transgenic plant comprises
 an altered trait selected from increased tolerance to abiotic stresses,
 increased tolerance to osmotic stress, increased tolerance to cold,
 increased germination in cold, increased tolerance to heat, increased
 germination in heat, increased tolerance to freezing conditions,
 increased tolerance to low nitrogen conditions, increased tolerance to
 low phosphate conditions, increased tolerance to disease, including
 fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 increased tolerance to multiple fungal pathogens, increased resistance to
 glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 increased sensitivity to ACC, altered sugar sensing, increased tolerance
 to sugars, altered carbon/nitrogen sensing, early flowering, late
 flowering, altered flower structure, loss of flower determinacy, reduced
 fertility, altered shoot meristem development, altered branching pattern,
 altered stem morphology, altered vascular tissue structure, reduced
 apical dominance, altered trichome density, altered trichome development,
 altered trichome structure, altered root development, altered shade
 avoidance, altered seed development, altered seed ripening, altered seed
 germination, slow growth, fast growth, altered cell differentiation,
 altered cell proliferation, altered cell expansion, altered phase change,
 altered senescence, abnormal embryo development, altered programmed cell
 death, lethality when overexpressed, altered necrosis patterns, increased
 plant size, increased biomass, large seedlings, dwarfed plants, dark
 green leaves, change in leaf shape, increased leaf size and mass, light
 green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 altered seed coloration, altered seed size, altered seed shape, large
 seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 content, altered seed protein content, altered seedprenyl content,
 altered leaf prenyl lipid content, increased anthocyanin levels, and
 decreased anthocyanin levels. Note: The sequence data for this patent did
 not form part of the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.
 Query Match 39.2%; Score 149.4; DB 12; Length 279;
 Best Local Similarity 74.0%; Pred. No. 1.7e-32;
 Matches 205; Conservative 0; Mismatches 66; Indels 6; Gaps 1;
 1 ATGGCGTCAAACTCAAGATTTCATCTCACATGAGCTTTAGTCAAAACAGATGTC 60
 1 ATGGCTCTAGTTCTATGAGCTCGAGCTTCTTGGAGCTCTAAGCAAAACAGATGTC 60
 61 GAGAGGGCTTTAGCGTTTACGATTAAGACACTCCCGACCGTTGGCAAAACAGATGTC 120
 61 GAGAGGGCTTTAGCGTTTACGATTAAGACACTCCCGACCGTTGGCAAAACAGATGTC 120
 121 GCTGTGCGAGGAAACATGTTAGAAAGTGAAGCCCACTATGACATTTCTCGTCAGAT 180
 121 GCAGTTGGAAGTAAATCTGCAGAGGAGTAAACGTCACGACATCTCGTTGAAGAT 180

QY 181 CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCCAATTACAAGACCTTGA-----A 234
 DB 181 CTCATGAACATCGAACAAGACTTAGTACCTTTGGCTTAATAACAACCCGTCGATGTGA 240
 QY 235 TCTAACTCAAGAAGCATCAATGACTTTTGACACAAGGT 271
 DB 241 AGTAAATCTAGAGGCATCGATGATTTTGATTGGAGT 277
 RESULT 4
 AAC34192
 ID AAC34192 standard; DNA; 478 BP.
 AC AAC34192;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5783.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 09-MAR-1999; 99US-0123180P.
 PR 23-MAR-1999; 99US-0123548P.
 PR 29-MAR-1999; 99US-0125788P.
 PR 29-MAR-1999; 99US-0126264P.
 PR 01-APR-1999; 99US-0126785P.
 PR 06-APR-1999; 99US-0127462P.
 PR 08-APR-1999; 99US-0128234P.
 PR 16-APR-1999; 99US-0128714P.
 PR 19-APR-1999; 99US-0129845P.
 PR 21-APR-1999; 99US-0130077P.
 PR 23-APR-1999; 99US-0130449P.
 PR 28-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 07-MAY-1999; 99US-0132486P.
 PR 11-MAY-1999; 99US-0132863P.
 PR 14-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 19-MAY-1999; 99US-0134768P.
 PR 20-MAY-1999; 99US-0134941P.
 PR 21-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.

Dd 2 AGCTCTTCTTGGACGCTTAAGCAAAACAAGATGTTTCGAGAGGCTTTAGCCGTTTACGAT 61
Qy 85 AAGGACACACCCGACCGATGGCAATGTGGCAAAAGCTGTGCGAGGGAACCTGTAGAA 144
Dd 62 AAAGACACTCCGACCGTGGCAAAAGCTGTGCGAGGTTGGAAGTAAATCTGCAGAG 121
Qy 145 GAAGTGAAGCGCCACTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGTGCTG 204
Dd 122 GAAGTTAAACGTCCTACTACGACATCTCTGTTGAAGATCTCATGAACATCGAACAAGACTTA 181
Qy 205 GTCCCTTTGCCCAATTACAGACCTTGA-----ATCTAACTCAAGAGCATCAATGAC 258
Dd 182 GTACCTTTGCTTAAATACAAACCGTCGATGTTTGAAGTAAATCTAGAGGAATCAATGAT 241
Qy 259 TTGACACAAAGGTATATACTAA 281
Dd 242 TTGATTTGAGGTATTGAAGAA 264

RESULT 5

ADW16587
ID ADW16587 standard; cDNA; 482 BP.
XX
AC ADW16587;
XX
DT 24-MAR-2005 (first entry)
XX
DE Eucalyptus grandis transcription factor cDNA MYB family Seq 306.

XX Gene; ss; plant; transcription; gene regulation; gene expression;
KW transgenic plant; drought resistance; disease resistance; salt tolerance;
KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
KW flower color.
XX
OS Eucalyptus grandis.
XX
PN WO2005001050-A2.
XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-US017965.
XX
PR 06-JUN-2003; 2003US-0476189P.
XX
PA (ARBO-) ARBORGEN LLC.

PI Bloksberg LN, Bryant C, Connott MB, Emerson SJ, Frost MJ;
PI Forster RL, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
DR WPI; 2005-075542/08.
DR P-PSDB; ADW17373.
XX
PT New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 3; SEQ ID NO 306; 1265pp; English.

XX This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA

CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polynucleotide is a
CC plant transcription factor cDNA sequence of the invention.

XX
SQ Sequence 482 BP; 137 A; 89 C; 121 G; 135 T; 0 U; 0 Other;

Query Match 32.2%; Score 122.8; DB 14; Length 482;
Best Local Similarity 70.9%; Pred. No. 8.5e-25;
Matches 163; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 28 TCACCATGACGCTTTAGTCAAAACAAGATGTTTCGAGAGGCTTGGCAGTTTACGACAAG 87
Dd 64 TCATCTGACGCCGGAAGCAGAACAAAGATGTTTCGAGAAGCGCATTTGCTCAATATGACAAG 123
Qy 88 GACACACCCGACCGATGGCAACATGTGGCAAAAGCTGTGCGAGGGAACCTGTAGAAAGAA 147
Dd 124 GACACTCCCGACAGGTGGCAGAAGATTGCCAAGGCCGTGGGGAATCTGCAGATGAA 183
Qy 148 GTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGTGCTGTC 207
Dd 184 GTGAAGACACTATGAATTTTATTCGAGGAGCTCAAGCACATCGAGTCTGGCAGATT 243
Qy 208 CCTTTGCCCAATTACAAGACCTTCAATCTAACTCAAGAAGCATCAATGA 257
Dd 244 CCTTTCTCTAATTACAGGTTCGAGCAACAATAGCGGCTGAAGGTTCGGTGA 293

RESULT 6

ADRS9489
ID ADRS9489 standard; cDNA; 760 BP.
XX
AC ADRS9489;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton cDNA sequence, SEQ ID 270.
XX
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
OS Gossypium hirsutum.
XX
PN US2004181830-A1.
XX
PD 16-SEP-2004.
XX
PF 29-JAN-2004; 2004US-00767795.
XX
PR 07-MAY-2001; 2001US-00849529.
PR 12-DEC-2001; 2001US-00021323.
XX
PA (KOVA/) KOVALIC D K.
PA (ZHOU/) ZHOU Y.
PA (CAOY/) CAO Y.
XX
PI Kovalic DK, Zhou Y, Cao Y;
XX WPI; 2004-667718/65.
XX
PT New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
PS Claim 1; SEQ ID NO 270; 14pp; English.

XX The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an

CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.

XX Sequence 760 BP; 308 A; 90 C; 144 G; 218 T; 0 U; 0 Other;

Query Match 31.7%; Score 120.6; DB 13; Length 760;
Best Local Similarity 75.4%; Pred. No. 4.3e-24;
Matches 150; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 32 CATGGACGTTTGTAGTCAAAACAAGATGTTCCGAGAGGGCTTGGCAGTTTACGACAAGGACA 91
DB 138 CATGGACAGCCAGCAAAACAAGATTTCCGAAGGGCTTTAGCTGTTACGACAAGGACA 197
QY 92 CACCCGACCGATGCGACAATGTGGCAAAAGCTGTCCGAGGGGAAACCTGTAGAGAAGTGA 151
DB 198 CACGAGATCGTTGGTACAATGTTGCTAAAGCTGTGGGAGGGAACCTGTTGAGGAAGTGA 257
QY 152 AGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGTCGTGTCCTT 211
DB 258 AGAAGCACTATGAGCTTCTTCTGGAAGATGTTAGACACATCGAGTCGGGTGCGGTTCCTT 317
QY 212 TGCCCAATTACAAGACCTT 230
DB 318 CCCCAGCTATTGGACCGT 336

RESULT 7

ACN61386/C
ID ACN61386 standard; cDNA; 543 BP.

XX ACN61386;

DT 02-DEC-2004 (first entry)

XX Cotton gynoecium tissue EST Clone ID: LIB3829-015-Q1-N6-E2, SEQ:16167.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
XX variety Nucotton338; library LIB3829; molecular tag; molecular marker;
XX genetic mapping; molecular mapping; seed germination; plant growth;
XX plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
PI WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.

XX Claim 1; SEQ ID NO 16167; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

XX Sequence 543 BP; 164 A; 128 C; 74 G; 177 T; 0 U; 0 Other;

Query Match 31.2%; Score 119; DB 13; Length 543;
Best Local Similarity 74.9%; Pred. No. 1.1e-23;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 32 CATGGACGTTTGTAGTCAAAACAAGATGTTCCGAGAGGGCTTGGCAGTTTACGACAAGGACA 91
DB 505 CATGGACAGCCAGCAAAACAAGATTTCCGAAGGGCTTTAGCTGTTTACGACAAGGACA 446
QY 92 CACCCGACCGATGCGACAATGTGGCAAAAGCTGTCCGAGGGGAAACCTGTAGAGAAGTGA 151
DB 445 CACGAGATCGTTGGTACAATGTTGCTTAAGCTGTGGAGAGAAAACCTGTTGAGGAAGTGA 386
QY 152 AGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGTCGTGTCCTT 211
DB 385 AGAAGCACTATGAGCTTCTTCTTGAAGATGTTAGACGATCGAGTCGGGTGCGGTTCCTT 326
QY 212 TGCCCAATTACAAGACCTT 230
DB 325 TCCCAGCTATTGGACCGT 307

RESULT 8
ADR59490

ADRS9490 standard; cDNA; 779 BP.
ADRS9490;
02-DEC-2004 (first entry)
Cotton cDNA sequence, SEQ ID 271.
Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
drought tolerance; plant disease resistance; galactomannan; lignin;
plant growth regulator; heat tolerance; herbicide tolerance;
homologous recombination; extreme osmotic condition tolerance;
pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
stress resistance.
Gossypium hirsutum.
US2004181830-A1.
16-SEP-2004.
29-JAN-2004; 2004US-00767795.
07-MAY-2001; 2001US-00849529.
12-DEC-2001; 2001US-00021323.
(KOVA/) KOVALIC D K.
(ZHOU/) ZHOU Y.
(CAOY/) CAO Y.
Kovalic DK, Zhou Y, Cao Y;
WPI; 2004-667718/65.
New recombinant nucleic acid molecules and polypeptides from Gossypium
hirsutum, useful for producing plants with improved biological
characteristics (e.g. improved plant cold or drought tolerance).
Claim 1; SEQ ID NO 271; 14pp; English.
The invention relates to a recombinant polynucleotide comprising any of
the 58798 Cotton plant cDNA sequences mentioned in the specification.
Also a recombinant polypeptide comprising any of the 58798 amino acid
sequences mentioned in the specification and producing a plant having an
improved property. Producing a plant having an improved property
comprises transforming a plant with a recombinant construct comprising a
promoter region functional in a plant cell operably joined to a
polynucleotide comprising a coding sequence for a polypeptide associated
with the property, and growing the transformed plant. The polypeptide is
useful for improving plant cold tolerance, manipulating growth rate in
plant cells by modification of the cell cycle pathway, improving plant
drought tolerance, providing increased resistance to plant disease,
producing galactomannan (or lignin or plant growth regulators), improving
plant heat tolerance, improving plant tolerance to herbicides, increasing
the rate of homologous recombination in plants, improving plant tolerance
to extreme osmotic conditions or to pathogens or pests, improving yield
by modification of photosynthesis, modifying seed oil or protein yield
and/or content, improving yield by modification of carbohydrate, nitrogen
or phosphorus use and/or uptake, or improving yield by providing improved
plant growth and development under at least one stress condition. The
polynucleotide and polypeptide may also be used in recombinant DNA
constructs, in physical arrays of molecules, as plant breeding markers,
or in computer-based storage and analysis systems. The present sequence
is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from USPTO at
seqdata.uspto.gov/sequences.html?DocID=20040181830. However only 6585
polynucleotide sequences were available, the remaining 52213
polynucleotides and all 58798 protein sequences were not present.
Sequence 779 BP; 248 A; 121 C; 157 G; 253 T; 0 U; 0 Other;
Query Match 31.2%; Score 119; DB 13; Length 779;

Best Local Similarity 74.9%; Pred. No. 1.3e-23;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 32 CATGGACGTTTGTAGTCAAAACAAGATGTTCCGAGAGGCGCTTGGCAGTTTACGACAAGACA 91
DB 172 CATGGACAGCCCAAGCAAAACAAAGATTTCCGAAAGGCGCTTTAGCTGTTTACGACAAGACA 231
QY 92 CACCCGACCGATGGCACAATGTGGCAAAAGCTGTGCGAGGGAACCTGTAGAGAAGTGA 151
DB 232 CACCAAGATCGTTGGTACATGTTGCTAAAGCTGTGGAGAGAAACCTGTTGAGAAAGTGA 291
QY 152 AGCGGCATCATGACATTTCTGCGAGGATCTCATCAACATCGAGACTGTCGTGTCCTT 211
DB 292 AGAAGCACTATGAGCTTCTTCTTGAAGATGTTAGAGCATCGAGTCGGTTCGGTTTCTT 351
QY 212 TCGCCCAATTACAAGACCTT 230
DB 352 TCCCGACTATTGGACCGT 370
RESULT 9
ADXS32012
ID ADXS32012 standard; cDNA; 704 BP.
XX AC ADXS32012;
XX DT 21-APR-2005 (first entry)
XX DE Plant full length insert polynucleotide seqid 14832.
XX KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PF 28-APR-2003; 2003US-00425114.
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 14832; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as

CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant calls by modification of the cell cycle pathway, for conferring
CC increased resistance to plant diseases, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX SQ Sequence 704 BP; 202 A; 159 C; 121 G; 222 T; 0 U; 0 Other;

Query Match 29.1%; Score 110.8; DB 13; Length 704;
Best Local Similarity 68.1%; Pred. No. 2.7e-21;
Matches 154; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCGTGGACGTTTGTAGTCAAAACAAGATGTTTC 60
DB 150 ATGGATCCAGTTCATCTCAGCTCTGGCTCATGGAGTGTAAAGCACACAGGCGCTTT 209
QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCGACCGATGGCACATGTGGCAAAA 120
DB 210 GAAAAGGCTTTAGCTGTTTATGACAGGACACTCTGACCGTTGGTACAATGTTGCTCAT 269
QY 121 GCTGTGGGAGGAAACTGTAGAAGAGTGAAGCGCCACTATGACATTCCTCGCAGGAT 180
DB 270 GCTGTTGGTGGCAAACTCCAGAGGAAGTGAAGGACACTACGAACTCTTGTTCAGGAT 329
QY 181 CTGATCAACATCGAGACTGTGTCCTCTTGGCCCAATTACAAGA 226
DB 330 GTTAAGCATATTGAGTCTGGACGTGTGCCATTCCCAAAATTACAAGA 375

RESULT 10
ADJ43786/c
ID ADJ43786 standard; cDNA; 466 BP.

XX AC ADJ43786;

XX DT 06-MAY-2004 (first entry)

XX DE Plant cDNA #4786.

XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX PD 22-JAN-2004.

XX PF 26-SEP-2002; 2002US-00260238.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 26-SEP-2001; 2001US-0325448P.

XX PR 04-APR-2002; 2002US-0370620P.

XX PA (BUDW/) BUDWORTH P.
XX PA (MOUG/) MOUGHAMER T.
XX PA (BRIG/) BRIGGS S P.
XX PA (COOP/) COOPER B.
XX PA (GLAZ/) GLAZEBROOK J.
XX PA (GOFF/) GOFF S A.
XX PA (KATA/) KATAGIRI F.

PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICHE D.
PA (ZHUT/) ZHU T.
XX
PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Riche D, Zhu T;
XX
XX WPI; 2004-190374/18.

XX New rice promoter, useful for manipulating crop plants to alter or
XX improve phenotypic characteristics, e.g. produce large quantities of oil
XX or proteins, resistance to insecticides, virus or fungi, stress tolerance
XX or high nutritional value.

XX Example 13; SEQ ID NO 4786; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,
XX leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
XX or constitutive transcription of an operatively linked nucleic acid
XX segment. The invention also relates to a method for augmenting a plant
XX genome and a method of identifying a gene, where its expression is
XX altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
XX in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
XX canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
XX sorghum, rice or wheat. The polynucleotides and the polypeptides they
XX encode are useful for manipulating crop plants to alter or improve
XX phenotypic characteristics, to produce large quantities of oil or
XX proteins, to incur resistance to insecticides, viruses or fungi, and to
XX incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
XX have a high nutritional value with reduced apical dominance or dwarfism,
XX early flowering or altered metabolic pathways. This sequence represents a
XX plant nucleic acid of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification but was obtained in
XX electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 466 BP; 93 A; 127 C; 124 G; 122 T; 0 U; 0 Other;

Query Match 28.9%; Score 110; DB 12; Length 466;
Best Local Similarity 67.4%; Pred. No. 3.8e-21;
Matches 155; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 25 ATCTCACCATGGACGTTTGTAGTCAAAACAAGATTTTCGAGAGGCGCTTGGCAGTTTACGAC 84
DB 407 AGCTCGTGTGGACCGCGGAGCAGAACAGATGTTTCGAGAAAGCCCTCGCGGTGACGAC 348

QY 85 AAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGAGGAGGAAACTGTAGAA 144
DB 347 AAGGACACGCGCCGACCGCTGGCACACGTGGCGCGCGCTCGCGGCAAGTCAGCGGAA 288

QY 145 GAAAGTGAAGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGTGCGT 204
DB 287 GAAGTGAAGCGGCACATGACGAGTGTGTTGGCGACATTATCTCATCGAGAGGCGGAC 228

QY 205 GTCCCTTTGCCCAATTACAGACCTTCGAATCTTAACCTCAAGACATCAA 254

DB 227 ATGCTCGAGCCCAATTACCGCTCTCCGCGCCACAGGGGATGAAGTGTCGA 178

RESULT 11
ACL30919/c
ID ACL30919 standard; cDNA; 513 BP.

XX AC ACL30919;

XX DT 02-JUN-2005 (first entry)

XX DE Rice abiotic stress response related polynucleotide SEQ ID NO:9482.
XX KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
XX KW agriculture.
XX OS Oryza sativa.

XX WO2003008540-A2.
XX 30-JAN-2003.
XX 21-JUN-2002; 2002WO-US019668.
XX 22-JUN-2001; 2001US-0300112P.
XX 24-AUG-2001; 2001US-0314662P.
XX 26-SEP-2001; 2001US-0325277P.
XX 21-NOV-2001; 2001US-0332132P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
XX Moughamer T, Provart N, Ricke D, Zhu T;
XX WPI; 2003-248011/24.
XX New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.
XX Disclosure; SEQ ID NO 9482; 89pp; English.
XX The invention relates to novel abiotic stress responsive polynucleotides
XX and polypeptides. Also disclosed are vectors, expression cassettes, host
XX cells, and plants containing such polynucleotides. Also disclosed are
XX methods for using the polynucleotides and polypeptides to alter the
XX responsiveness of a plant to abiotic stress. The invention is useful in
XX agriculture. The nucleic acid is useful for determining whether a test
XX plant has been exposed to an abiotic stress condition. It is also useful
XX for selecting an agent that alters abiotic stress regulated
XX polynucleotide expression in a plant cell, and to identify a homolog or
XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX molecule and the polypeptide encoded by it are useful in altering the
XX responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX stress, osmotic stress or any of their combinations. The present sequence
XX is used in the exemplification of the invention
SQ Sequence 513 BP; 126 A; 128 C; 127 G; 131 T; 0 U; 1 Other;
Query Match 28.5%; Score 108.6; DB 11; Length 513;
Best Local Similarity 70.9%; Pred. No. 1e-20;
Matches 144; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 27 CTCACATGGAGCTTTAGTCAAAACAAGATGTCGAGAGGGCCTTGGCAGTTTACGACA 86
Db 476 CTCCTCTGGACCGGAGGAGACAAGATGTCGAGAGGCACTGGCGGTGTACGACAG 417
QY 87 GGACACACCCGACCGATGCAATGTCGCAAAAGCTGTCGAGGGAACTGTAGAAGA 146
Db 416 GGACACCCCGACCGCTGTCGCAAGATCGCCGGGCCATCGCGGGAAGACCGCGGACGA 357
QY 147 AGTGAAGCGCCACTATGACATTCCTGCTCAGGATCTCATCAACATCGAGACTGGTTCGTGT 206
Db 356 AGTGAAGCGCTACTACGACCTGCTGTCGAGAGCGTGGCCGCATCGAGGCTGGCCAAAT 297
QY 207 CCCTTTGCCCAATTAACAAGCT 229
Db 296 GCGGTACGCCAATTAACCGTCTCT 274
RESULT 12
AAC35256
ID AAC35256 standard; DNA; 508 BP.
XX AAC35256;
AC AAC35256;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 9536.
XX

KW Hybridisation assay; Genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway; metabolic pathway;
KW Promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 23-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131445P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132488P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015151438P.

PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 23-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 28.5%; Score 108.4; DB 3; Length 508;

Best Local Similarity 68.0%; Pred. No. 1.1e-20; Mismatches 0; Gaps 0;
Matches 151; Conservative 0; Indels 71; Delele 0;

Qy 1 ATGGCGTCAAACCTCAAGAAGTTCAATCTCACCATGACAGCTTTAGTCAAAACAAGATGTTTC 60
|||||
Db 108 ATGGCTTCCAACCTCAATGAGCTCTAGCGCTTCTTGACACGCTAAGGAGAAACAAATATTAT 167
|||||
Qy 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGCGATGGCACATGTGGCAAAA 120
|||||
Db 168 GAAAGGGCGTGGCTACATATGACCAGGACACTCTCGACCGTTGGCATACGTTGCAAGA 227
|||||
Qy 121 GCTGTGCGAGGAAAACTGTAGAAGAACTGTAAGCGGCACCTATGACATTTCTCGTCAGGAT 180
|||||
Db 228 GCCGTTGGCGGCAATCAGCTGAAGAAGTAAGCGGACACTACGAGCTCTCTATTAGGAT 287
|||||
Qy 181 CTCATCAACATCGAGACTGGTGTGTCCTTTTCCCAATTAC 222
|||||
Db 288 GTCAATGACATTGAGTCAGGCGCTTATCCACATCCCAATTAC 329
|||||

RESULT 13
ACN59055
ID ACN59055 standard; cDNA; 577 BP.
XX
AC ACN59055;

CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nu cotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

XX
SQ Sequence 505 BP; 131 A; 114 C; 102 G; 158 T; 0 U; 0 Other;

Query Match 28.2%; Score 107.4; DB 13; Length 505;
Best Local Similarity 70.2%; Pred. No. 2.2e-20;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 TCAATCTCACCATGGACGTTTGTCTCAAAACAAGATGTCGAGAGGCGCTTGGCAGTTTAC 81
Db 466 TCAGGTTTCAAGTGGACGCCAAGCAAAACAAGTTGTTTCGAGAACGCTTTAGTCTATCTAC 407
Qy 82 GACAAGGACACACCCGACCGCATGCGCAATGTGGCAAAAGCTGTGCGAGGGGAAACTGTA 141
Db 406 GACAGGATACACAGCCGCTGGCACAAGCTAGCCAGGCGCTGTGGGGGAGACCGTG 347
Qy 142 GAAGAAGTGAAGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGGT 201
Db 346 GAGGAAGTGAAGTTGATTACCAAGACCTTTGTGATGACATCAAGCAGATAGAGTCTGGG 287
Qy 202 CGTGTCCTTTGCCANTTACAAGA 226
Db 286 CACGTGCTTTGGCCCCCTTACAAGA 262

RESULT 15

ACN59701
ID ACN59701 standard; cDNA; 532 BP.

XX
AC ACN59701;

DT 02-DEC-2004 (first entry)

XX
DE Cotton gynoecium tissue EST Clone ID: LIB3829-034-Q1-K6-A10, SEQ:14482.

XX
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW variety Nu cotton33B; library LIB3829; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.

XX
OS Gossypium hirsutum.

XX
FN US2004123340-A1.

XX
PD 24-JUN-2004.

XX
PF 12-DEC-2001; 2001US-00021323.

XX
PR 14-DEC-2000; 2000US-0255619P.

XX
PA (DEIK/) DEIKMAN J.

XX
PA (FENG/) FENG P C C.

XX
PA (FING/) FINCHER K L.

XX
PA (ZIEG/) ZIEGLER T E.

XX
PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX
DR WPI; 2004-479808/45.

XX

PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.

XX
PS Claim 1; SEQ ID NO 14482; 34pp; English.

XX
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nu cotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nu cotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

XX
SQ Sequence 532 BP; 170 A; 104 C; 119 G; 139 T; 0 U; 0 Other;

Query Match 28.2%; Score 107.4; DB 13; Length 532;
Best Local Similarity 70.2%; Pred. No. 2.2e-20;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 TCAATCTCACCATGGACGTTTGTCTCAAAACAAGATGTCGAGAGGCGCTTGGCAGTTTAC 81
Db 56 TCAGGTTTCAAGTGGACGCCAAGCAAAACAAGTTGTTTCGAGAACGCTTTAGTCTATCTAC 115
Qy 82 GACAAGGACACACCCGACCGCATGCGCAATGTGGCAAAAGCTGTGCGAGGGGAAACTGTA 141
Db 116 GACAAGGATACACAGCCGCTGGCACAAGCTAGCCAGGCGCTGTGGGGGAGACCGTG 175
Qy 142 GAAGAAGTGAAGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGGT 201
Db 176 GAGGAAGTGAAGTTGATTACCAAGACCTTTGTGATGACATCAAGCAGATAGAGTCTGGG 235
Qy 202 CGTGTCCTTTGCCAAATTACAAGA 226
Db 236 CACGTGCTTTGGCCCCCTTACAAGA 260

Search completed: March 10, 2006, 18:43:29
Job time : 404 Secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:35:45 ; Search time 442 Seconds
(without alignments)
657.736 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNSRSSIPMTFSQKMF.....HSSDFEKFQKLVSYISLV 126

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10697787/runat_09032006_150641_3661/app_query.fasta 1
-DB=Published Applications NA New -QPM=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs05p
-USER=US10697787 @CGN_1_1_653 @runat_09032006_150641_3661 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_New.*

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description |
|------------|-------|-------------|--------|---------------------------------------|
| 1 | 269.5 | 41.1 | 517 9 | US-11-096-568A-8355 Sequence 8355, Ap |
| 2 | 214 | 32.6 | 1724 9 | US-11-096-568A-4492 Sequence 4492, Ap |
| 3 | 211 | 33.2 | 866 7 | US-10-714-887-245 Sequence 245, App |
| 4 | 204 | 31.1 | 896 7 | US-10-714-887-273 Sequence 273, App |

| | | | | | | |
|------|-------|------|--------|----|-----------------------|--------------------|
| 5 | 193 | 29.4 | 957 | 9 | US-11-087-099-102 | Sequence 102, App |
| 6 | 187 | 28.5 | 876 | 9 | US-11-096-568A-29719 | Sequence 29719, A |
| 7 | 185.5 | 28.3 | 1254 | 9 | US-11-096-568A-21266 | Sequence 21266, A |
| 8 | 151.5 | 23.1 | 994 | 9 | US-11-096-568A-23894 | Sequence 23894, A |
| 9 | 146 | 22.3 | 1351 | 9 | US-11-096-568A-20386 | Sequence 20386, A |
| 10 | 126 | 19.2 | 1431 | 9 | US-11-096-568A-13836 | Sequence 13836, A |
| 11 | 124 | 18.9 | 933 | 9 | US-11-096-568A-24021 | Sequence 24021, A |
| 12 | 121 | 18.4 | 600 | 12 | US-11-136-527-4714 | Sequence 4714, Ap |
| 13 | 121 | 18.4 | 2119 | 12 | US-11-136-527-618 | Sequence 618, Ap |
| C 14 | 96 | 14.6 | 1593 | 8 | US-10-750-185-30681 | Sequence 30681, A |
| C 15 | 96 | 14.6 | 1593 | 8 | US-10-750-623-30681 | Sequence 30681, A |
| C 16 | 83.5 | 12.7 | 607 | 6 | US-09-925-065A-206387 | Sequence 206387, A |
| C 17 | 83.5 | 12.7 | 607 | 6 | US-09-925-065A-206388 | Sequence 206388, A |
| 18 | 83.5 | 12.7 | 2919 | 7 | US-10-932-182A-3711 | Sequence 3711, Ap |
| 19 | 83.5 | 12.7 | 2919 | 7 | US-10-932-182A-3711 | Sequence 3711, Ap |
| 20 | 82 | 12.5 | 1506 | 9 | US-11-096-568A-22860 | Sequence 22860, A |
| C 21 | 80.5 | 12.3 | 607 | 6 | US-09-925-065A-206386 | Sequence 206386, A |
| C 22 | 79 | 12.0 | 499 | 9 | US-11-096-568A-14174 | Sequence 14174, A |
| C 23 | 78.5 | 12.0 | 607 | 6 | US-09-925-065A-140607 | Sequence 140607, A |
| C 24 | 78.5 | 12.0 | 162173 | 12 | US-11-121-086-72 | Sequence 72, Appl |
| 25 | 75.5 | 11.5 | 1509 | 9 | US-11-096-568A-26086 | Sequence 26086, A |
| 26 | 75.5 | 11.5 | 1545 | 9 | US-11-096-568A-9922 | Sequence 9922, Ap |
| 27 | 75 | 11.4 | 1332 | 9 | US-11-096-568A-33483 | Sequence 33483, A |
| 28 | 75 | 11.4 | 3276 | 7 | US-10-932-182A-190881 | Sequence 190881, A |
| 29 | 75 | 11.4 | 3276 | 7 | US-10-932-182A-190881 | Sequence 190881, A |
| 30 | 74.5 | 11.4 | 1779 | 7 | US-10-932-182A-79555 | Sequence 79555, A |
| 31 | 74.5 | 11.4 | 1779 | 7 | US-10-932-182A-79555 | Sequence 79555, A |
| 32 | 74 | 11.3 | 885 | 7 | US-10-932-182A-78621 | Sequence 78621, A |
| 33 | 74 | 11.3 | 885 | 7 | US-10-932-182A-78621 | Sequence 78621, A |
| 34 | 74 | 11.3 | 2550 | 7 | US-10-932-182A-77658 | Sequence 77658, A |
| 35 | 74 | 11.3 | 2550 | 7 | US-10-932-182A-77658 | Sequence 77658, A |
| 36 | 73 | 11.1 | 864 | 9 | US-11-096-568A-28818 | Sequence 28818, A |
| 37 | 72.5 | 11.1 | 1230 | 7 | US-10-932-182A-82655 | Sequence 82655, A |
| 38 | 72.5 | 11.1 | 1230 | 7 | US-10-932-182A-82655 | Sequence 82655, A |
| 39 | 71.5 | 10.9 | 2295 | 12 | US-11-055-557-19 | Sequence 19, Appl |
| 40 | 71.5 | 10.9 | 2307 | 12 | US-11-055-557-3 | Sequence 3, Appl |
| 41 | 71.5 | 10.9 | 2307 | 12 | US-11-055-557-5 | Sequence 5, Appl |
| 42 | 71.5 | 10.9 | 6425 | 12 | US-11-131-479-90 | Sequence 90, Appl |
| 43 | 71 | 10.8 | 1638 | 7 | US-10-932-182A-79707 | Sequence 79707, A |
| 44 | 71 | 10.8 | 1638 | 7 | US-10-932-182A-79707 | Sequence 79707, A |
| 45 | 70.5 | 10.7 | 1390 | 9 | US-11-096-568A-5907 | Sequence 5907, Ap |

ALIGNMENTS

RESULT 1

US-11-096-568A-8355
; Sequence 8355, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8355
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(517)
; OTHER INFORMATION: Ceres Seq. ID no. 15225172
; US-11-096-568A-8355

| | | | |
|------------------------|----------|---------------|-----|
| Alignment Scores: | 6.99e-25 | Length: | 517 |
| Pred. No.: | 269.50 | Matches: | 55 |
| Score: | 66.7% | Conservative: | 13 |
| Best Local Similarity: | 53.9% | Mismatches: | 31 |
| Query Match: | 41.1% | Indels: | 3 |

| | | | |
|--|----------|--|------|
| DB: | 9 | Gaps: | 1 |
| US-10-697-787-2 (1-126) x US-11-096-568A-8355 (1-517) | | | |
| Qy | 3 | SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg | 22 |
| Db | 63 | GCTACAAACATGGCCCTCAAGTCAGGTTGGACTCCGAAGCAACAGAGATTGGAGAT | 122 |
| Qy | 23 | AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal | 42 |
| Db | 123 | GCCTTGGCATCTTCGACAGGACACCCACAGAGTGGCACACGGTGGCCAGGCGGTC | 182 |
| Qy | 43 | GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle | 62 |
| Db | 183 | GGAGGAAAAACGGTGGAGGAAGTGAAAGGCATTATGAGAAGCTCGTGGAAAGATGGAAG | 242 |
| Qy | 63 | AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg | 82 |
| Db | 243 | GAGATTGAGGAAGGTACGTGCCCTCCCAATTACCGA-----AGTGTGCAAGA | 293 |
| Qy | 83 | SerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIle | 102 |
| Db | 294 | GCCACGGTTACATGGATGAAGAAACAGGATGAGGTTCTGAGTCTGCAGTGAAGTGTG | 353 |
| Qy | 103 | TyrPhe 104 | |
| Db | 354 | ATTAT 359 | |
| RESULT 2 | | | |
| US-11-096-568A-4492 | | | |
| ; Sequence 4492, Application US/11096568A | | | |
| ; Publication No. US20060048240A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Alexandrov, Nikolai et al. | | | |
| ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides | | | |
| ; FILE REFERENCE: 2750-1592PUS2 | | | |
| ; CURRENT APPLICATION NUMBER: US/11/096,568A | | | |
| ; CURRENT FILING DATE: 2005-04-01 | | | |
| ; NUMBER OF SEQ ID NOS: 34471 | | | |
| ; SEQ ID NO 4492 | | | |
| ; LENGTH: 1724 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Glycine max | | | |
| ; FEATURE: | | | |
| ; NAME/KEY: misc feature | | | |
| ; LOCATION: (1)..(1724) | | | |
| ; OTHER INFORMATION: Ceres Seq. ID no. 13638441 | | | |
| US-11-096-568A-4492 | | | |
| Alignment Scores: | | | |
| Pred. No.: | 4.54e-17 | Length: | 1724 |
| Score: | 214.00 | Matches: | 41 |
| Percent Similarity: | 71.8% | Conservative: | 12 |
| Best Local Similarity: | 55.4% | Mismatches: | 19 |
| Query Match: | 32.6% | Indels: | 2 |
| DB: | 9 | Gaps: | 1 |
| US-10-697-787-2 (1-126) x US-11-096-568A-4492 (1-1724) | | | |
| Qy | 3 | SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg | 22 |
| Db | 560 | AATAATAGCAGGAGC-----ACAAAATGGACCTCTGAAGAGAACAAAGCTCTTTGAAAT | 613 |
| Qy | 23 | AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal | 42 |
| Db | 614 | GCTCTTGAGTGCATGATAGACACCCCGGATCGGTGGCACAGTGGCTGAGATGATT | 673 |
| Qy | 43 | GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle | 62 |
| Db | 674 | CCTGGAAGACAGTGGTTGATGTGATGAAGGCAGTACAAGGAGTTGGAAGTAGATGTTAGC | 733 |
| Qy | 63 | AsnIleGluThrGlyArgValProLeuProAsnTyrLysThr | 76 |

| | | | |
|---|---------|--|-----|
| Db | 734 | AATATAGAAGCTGGTTTGATTCCAGTTCCTCGCTATAGTAGC | 775 |
| RESULT 3 | | | |
| US-10-714-887-245 | | | |
| ; Sequence 245, Application US/10714887 | | | |
| ; Publication No. US20060015972A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Mendel Biotechnology, Inc. | | | |
| ; APPLICANT: HEARD, Jacqueline | | | |
| ; APPLICANT: RIECHMANN, Jose Luis | | | |
| ; APPLICANT: CREELMAN, Robert | | | |
| ; APPLICANT: RATCLIFFE, Oliver | | | |
| ; APPLICANT: CANALES, Roger | | | |
| ; APPLICANT: REPETTI, Peter | | | |
| ; APPLICANT: KUMIMOTO, Roderick W | | | |
| ; APPLICANT: GUTTERSON, Neal | | | |
| ; APPLICANT: REUBER, T. Lynne | | | |
| ; APPLICANT: PINEDA, Omaika | | | |
| ; APPLICANT: SHERMAN, Bradley K | | | |
| ; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS | | | |
| ; FILE REFERENCE: MBI0058-CIP | | | |
| ; CURRENT APPLICATION NUMBER: US/10/714,887 | | | |
| ; CURRENT FILING DATE: 2003-11-13 | | | |
| ; PRIOR APPLICATION NUMBER: 10/412,699 | | | |
| ; PRIOR FILING DATE: 2003-04-10 | | | |
| ; PRIOR APPLICATION NUMBER: 09/506,720 | | | |
| ; PRIOR FILING DATE: 2000-02-17 | | | |
| ; PRIOR APPLICATION NUMBER: 60/135,134 | | | |
| ; PRIOR FILING DATE: 1999-05-20 | | | |
| ; PRIOR APPLICATION NUMBER: 09/394,519 | | | |
| ; PRIOR FILING DATE: 1999-09-13 | | | |
| ; PRIOR APPLICATION NUMBER: 09/533,392 | | | |
| ; PRIOR FILING DATE: 2000-03-22 | | | |
| ; PRIOR APPLICATION NUMBER: 09/533,029 | | | |
| ; PRIOR FILING DATE: 2000-03-22 | | | |
| ; PRIOR APPLICATION NUMBER: 09/532,591 | | | |
| ; PRIOR FILING DATE: 2000-03-22 | | | |
| ; PRIOR APPLICATION NUMBER: 09/533,030 | | | |
| ; PRIOR FILING DATE: 2000-03-22 | | | |
| ; PRIOR APPLICATION NUMBER: 60/125,814 | | | |
| ; PRIOR FILING DATE: 1999-03-23 | | | |
| ; PRIOR APPLICATION NUMBER: 09/713,994 | | | |
| ; PRIOR FILING DATE: 2000-11-16 | | | |
| ; Remaining Prior Application data removed - See File Wrapper or PALM. | | | |
| ; NUMBER OF SEQ ID NOS: 430 | | | |
| ; SOFTWARE: PatentIn version 3.2 | | | |
| ; SEQ ID NO 245 | | | |
| ; LENGTH: 866 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Arabidopsis thaliana | | | |
| ; FEATURE: | | | |
| ; OTHER INFORMATION: G2701 reference sequence; predicted polypeptide sequence is para | | | |
| US-10-714-887-245 | | | |
| Alignment Scores: | | | |
| Pred. No.: | 4.6e-17 | Length: | 866 |
| Score: | 211.00 | Matches: | 41 |
| Percent Similarity: | 64.6% | Conservative: | 10 |
| Best Local Similarity: | 51.9% | Mismatches: | 28 |
| Query Match: | 32.2% | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |
| US-10-697-787-2 (1-126) x US-10-714-887-245 (1-866) | | | |
| Qy | 1 | MetAlaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe | 20 |
| Db | 109 | ATGGTAGCTTACACAGCTCGAGTACGGGTAGCTGGACTAAAGAGAGACAAGATGTC | 168 |
| Qy | 21 | GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys | 40 |
| Db | 169 | GAACGAGCTCTTCGATATACGCTGAAGACTCGCTGATCGCTGGTTTAAAGTTGCTTCC | 228 |

QY 41 AlavalGlyLysThrValGluGluValLysArgHisThrAspLeuValGluAsp 60
Db 229 ATGATCCTCGAAGACTGTTTGGTATGATCAAGCAATATAGTAAGCTTTGAAGAGAC 288
QY 61 LeuileAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSer 79
Db 289 GTTTCGATTAATGAAGAGAGCGTGTGCCCATCTCTGGTTATCCTCGACGCTTCTCT 345

RESULT 4

US-10-714-887-273
; Sequence 273, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: FINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 273
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:

US-10-714-887-273
; OTHER INFORMATION: G1634; predicted polypeptide sequence is paralogous to G2701

Alignment Scores:
Pred. No.: 3,83e-16 Length: 896
Score: 204.00 Matches: 38
Percent Similarity: 68.7% Conservative: 8
Best Local Similarity: 56.7% Mismatches: 21
Query Match: 31.1% Indels: 0
DB: 7 Gaps: 0

US-10-697-787-2 (1-126) x US-10-714-887-273 (1-896)

QY 10 SerProTrrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLys 29
Db 106 TCGAGCTGGACTTAAGAAGACCAAGAAGTTTGAGCGAGCTCTTGCTGTCTACGCTGAT 165

QY 30 AspThrProAspArgTrrHisnValAlaLysAlaValGlyLysThrValGluGlu 49
Db 166 GACAGCCCTGATCGCTGGTTCAAAGTTGCTGTATGATCCCTGGAAAGACCATATCAGAT 225
QY 50 ValLysArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgVal 69
Db 226 GTCATGAGGCAATACTCTAAGCTTGAAGAAGACCTCTTCGATATCGAAGCAGGACTTGTC 285
QY 70 ProLeuProAsnTyrLysThr 76
Db 286 CCGATCCCGGGTTACCGTTCA 306

RESULT 5

US-11-087-099-102
; Sequence 102, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 102
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Glycine max
; US-11-087-099-102

Alignment Scores:

Pred. No.: 1,09e-14 Length: 957
Score: 193.00 Matches: 43
Percent Similarity: 48.5% Conservative: 7
Best Local Similarity: 41.7% Mismatches: 25
Query Match: 29.4% Indels: 28
DB: 9 Gaps: 2

US-10-697-787-2 (1-126) x US-11-087-099-102 (1-957)

QY 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 109 TGGACTAGAGAAGATAACAAGAAGTTTGAAGTGCCCTTGCTATATATGACAAGGACACC 168
QY 32 ProAspArgTrrHisnValAlaLysAlaValGlyLysThrValGluGluValLys 51
Db 169 CCAGATAGATGTTCAAGGTGGCCGCTATGATCCCTGGGAAGACTGTGTTTGATGTGATC 228
QY 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
Db 229 AAGCAATATAGGAAGACTCGAAGAGAGATGTGAGTGAATCGAAGACGAGCATGTTCCGATT 288
QY 72 ProAsnTyrLys-----ThrPheGluSerAsnSerArgSerIleAsnAspPhe 87
Db 289 CCGGCTACCTTGCACTCTCTTTCACCTTTGAG----- 321
QY 88 AspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIleTyrPheAspAsnHis 107
Db 322 -----CTGTTTGACAACACCAC 336

RESULT 6

US-11-096-568A-29719
; Sequence 29719, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1592PUS2

```
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29719
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(876)
; OTHER INFORMATION: Ceres Seq. ID no. 4936979
US-11-096-568A-29719

Alignment Scores:
Pred. No.: 5.76e-14 Length: 876
Score: 187.00 Matches: 35
Percent Similarity: 68.3% Conservative: 8
Best Local Similarity: 55.6% Mismatches: 20
Query Match: 28.5% Indels: 0
DB: 9 Gaps: 0

US-10-697-787-2 (1-126) x US-11-096-568A-29719 (1-876)

Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 64 TGGACAGCTGAAGAAAACAAGAAATTCGAAACGCITTTAGCCTTTTACGACAAAGATACT 123

Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
Db 124 CCCGACAGATGGTCCACAGTCGCTGCCATCTTCCCGTAAACAGTCGGAGATGTGATC 183

Qy 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
Db 184 AAACAATACAGAGAGCTTGAGGAAGACGTTAGCGACATCGAAGCTGCTTTATACCAATC 243

Qy 72 ProAsnTyr 74
Db 244 CTTGGTTAC 252

RESULT 7
US-11-096-568A-21266
; Sequence 21266, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21266
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1254)
; OTHER INFORMATION: Ceres Seq. ID no. 12402516
US-11-096-568A-21266

Alignment Scores:
Pred. No.: 1.42e-13 Length: 1254
Score: 185.50 Matches: 41
Percent Similarity: 61.4% Conservative: 10
Best Local Similarity: 49.4% Mismatches: 23
Query Match: 28.3% Indels: 2
DB: 9 Gaps: 2

US-10-697-787-2 (1-126) x US-11-096-568A-21266 (1-1254)

Qy 1 MetAlaSerAsnSerArgSerSer-----IleSerProTrpThr 13
||||| ||| |||||
```

```
Db 151 ATGGCCGCTGAACGAGCGCAGCAGCTGCGCGTGGGGAGGCGGTGCGGTGCTGGACC 210
Qy 14 PheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyr-----AspLysAspThr 31
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 CGCGAGCAGGAGAGCGCTTCGAGAACGCGGTGCGCAGCATGGCGCGGAGGAGGACGG 270
Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GACGCGAGGTGGAGAAAGTCAGCGAGCGCGTCGAGGGGAAGACGCCGCGGAGGAGTCA 330
Qy 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 CGGCACTACGAGCTGCTGCTGGAGGACGTCGAGCGCATCGAGTCGGGCGCGCTCCGCT 390

Qy 72 ProAsnTyr 74
||| |||
Db 391 CCGGCGGTAC 399

RESULT 8
US-11-096-568A-23894
; Sequence 23894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23894
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(994)
; OTHER INFORMATION: Ceres Seq. ID no. 12416292
US-11-096-568A-23894

Alignment Scores:
Pred. No.: 2.52e-09 Length: 994
Score: 151.50 Matches: 30
Percent Similarity: 64.1% Conservative: 11
Best Local Similarity: 46.9% Mismatches: 22
Query Match: 23.1% Indels: 1
DB: 9 Gaps: 1

US-10-697-787-2 (1-126) x US-11-096-568A-23894 (1-994)

Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 TGCACGCGGAGGAGAACAAAGCTGTTTCGAGAAAGCAGCTGCGCAGATCGCACCGAAC 263

Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CCCGACAGGTGGGAGAAAGTGGCGCGGTCTCCCTCGGAAGACGTCGACGACGTGAGG 323

Qy 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThr---GlyArgValPro 70
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 AGCCAGTACCACGCCCTCGAGAGAGGACGTGGGCGTCTATCGAGGCGCGGCTCGTCC 383

Qy 71 LeuProAsnTyr 74
||| |||
Db 384 TTCGCGGTAC 395

RESULT 9
US-11-096-568A-20386
; Sequence 20386, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
```



```
US-11-136-527-4714
; Sequence 4714, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4714
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4714

Alignment Scores:
Pred. No.: 1.13e-05 Length: 600
Score: 121.00 Matches: 22
Percent Similarity: 61.2% Conservative: 8
Best Local Similarity: 44.9% Mismatches: 19
Query Match: 18.4% Indels: 0
DB: 12 Gaps: 0

US-10-697-787-2 (1-126) x US-11-136-527-4714 (1-600)

Qy 11 ProTrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAsp 30
Db 308 CCWTGGACAACAGAGRGCGAAACTTGTAGAACAGGCTTTGAARACATACCCGGTAAT 367
Qy 31 ThrProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluVal 50
Db 368 ACTCCGGAAGATGGGAAAGATCGCGAAGCGCTTCTYGGCAGGACGAAGAGGACTGC 427
Qy 51 LysArgHisTyrAspIleLeuValGlu 59
Db 428 ATGAGCGCATACAGGAGCTCGTGGAG 454

RESULT 13
US-11-136-527-618
; Sequence 618, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 618
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-618

Alignment Scores:
Pred. No.: 5.59e-05 Length: 2119
Score: 121.00 Matches: 22
Percent Similarity: 61.2% Conservative: 8
Best Local Similarity: 44.9% Mismatches: 19
Query Match: 18.4% Indels: 0
DB: 12 Gaps: 0

US-10-697-787-2 (1-126) x US-11-136-527-618 (1-2119)

Qy 11 ProTrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAsp 30
Db 1827 CCWTGGACAACAGAGRGCGAAACTTGTAGAACAGGCTTTGAARACATACCCGGTAAT 1886
Qy 31 ThrProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluVal 50
Db 1887 ACTCCGGAAGATGGGAAAGATCGCGAAGCGCTTCTYGGCAGGACGAAGAGGACTGC 1946
Qy 51 LysArgHisTyrAspIleLeuValGlu 59
Db 1947 ATGAGCGCATACAGGAGCTCGTGGAG 1973

RESULT 14
US-10-750-185-30681/c
; Sequence 30681, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30681
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Bovine 19866881746860
US-10-750-185-30681

Alignment Scores:
Pred. No.: 0.0645 Length: 1593
Score: 96.00 Matches: 31
Percent Similarity: 51.4% Conservative: 23
Best Local Similarity: 29.5% Mismatches: 33
Query Match: 14.6% Indels: 20
DB: 8 Gaps: 5

US-10-697-787-2 (1-126) x US-10-750-185-30681 (1-1593)

Qy 16 GlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTyr 35
Db 1591 GAACAGAGCTTTTGGAAACAGCTTTGAAACATACCCAGTAATACACCTGAAAGATGG 1532
Qy 36 HisAsnValAlaLysAlaValGlyGlyLysThrValGluValLysArgHisTyrAsp 55
Db 1531 GAAAAAATAGCAGACAGCAGTGCCTGGCAGCACAAGAGGACTGCATGAACGATACAAG 1472
Qy 56 IleLeu-----ValGluAspLeuIleAsnIleGluThrGlyArgVal----- 69
Db 1471 GTTGTAGATCTTGAGTTTATAGATTCATATCAT-TTAAATTATGCTGTATACTTACT 1413
Qy 70 ProLeuPro-----AsnTyrLysThrPheGluSerAsnSer----- 81
Db 1412 TCAATACCAACCATAGTTGGGGCCAGTATTAGACATTTCTAGCTCGCAATTAATTT 1353
Qy 82 ArgSerIleAsnAspPheAspThrArgTyrIleThr-----LysTyrLeuTyr 97
Db 1352 AAAAAATCAACCT-TTTCAGTACAGATTTATCCAATTTGAATGTTTACATATTTATAT 1294
Qy 98 MetMetLeuSerIle 102
Db 1293 CAAACACTAGAAATG 1279

RESULT 15
```

US-10-750-623-30681/c
; Sequence 30681, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750, 623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30681
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Bovine 19866881746860
US-10-750-623-30681

Alignment Scores:
Pred. No.: 0.0645 Length: 1593
Score: 96.00 Matches: 31
Percent Similarity: 51.4% Conservative: 23
Best Local Similarity: 29.5% Mismatches: 33
Query Match: 14.6% Indels: 20
DB: 8 Gaps: 5

US-10-697-787-2 (1-126) x US-10-750-623-30681 (1-1593)

```
QY 16 GlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrp 35
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1591 GAACAGAGCTTTTGGACAAGCTTTGAAAACATACCCAGTAAATACACCTGAAAGATGG 1532

QY 36 HisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAsp 55
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1531 GAAAAAATAGCAGAGAGCAGTGCCTGGCAGGCAAGAGAGGAGGCTGCATGAAACGATACAAG 1472

QY 56 IleLeu-----ValGluAspLeuIleAsnIleGluThrGlyArgVal----- 69
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1471 GTTGTAGATCTTGAGTTTATAGATCCATTATCAT-TTAAATTATGGTGTATACTTACT 1413

QY 70 ProLeuPro-----AsnTyrLysThrPheGluSerAsnSer----- 81
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1412 TCATATACCACCATAGTTGTGGGCCAGTATTAGACATTTCTAGCTCTGGCAATTAATTT 1353

QY 82 ArgSerIleAsnAspPheAspThrArgTyrIleThr-----LysTyrLeuTyr 97
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1352 AAAAAAATCAACCT-TTTCAGTACAGATTTATCCAATTTGAATGTTTACATATTTATAT 1294

QY 98 MetMetLeuSerIle 102
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1293 CAAACACTAGAAATG 1279
```

Search completed: March 10, 2006, 19:45:45
Job time : 445 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:35:01 ; Search time 657 Seconds
(without alignments)
1585.908 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNGRSSISPTWFSQKMF.....HSSDFEKFQKLVSVISIV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134685005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/abs9/ABSSWEB spool/US10697787/runat_09032006_150638_3580/app_query.fasta_1
-DB=Published_Applications NA_Main -QFWT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -NORM=score=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10697787 @CGN 1 1 1025 @runat_09032006_150638_3580 -NCPU=6 -ICPU=3
-NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------------------|-------------------|
| 1 | 656 | 100.0 | 381 | 8 | US-10-697-787-1 | Sequence 1, Appli |
| 2 | 656 | 100.0 | 381 | 9 | US-10-512-600-1 | Sequence 1, Appli |
| 3 | 293 | 44.7 | 678 | 7 | US-10-424-599-6644 | Sequence 6644, Ap |
| 4 | 290 | 44.2 | 760 | 8 | US-10-767-795-270 | Sequence 270, App |
| 5 | 287 | 43.8 | 704 | 7 | US-10-425-114-14832 | Sequence 14832, A |
| 6 | 287 | 43.8 | 730 | 7 | US-10-424-599-135729 | Sequence 135729, |
| 7 | 286 | 43.6 | 556 | 3 | US-09-770-152-473 | Sequence 473, App |

| | | | | | | |
|------|-------|------|------|---|----------------------|-------------------|
| 8 | 286 | 43.6 | 578 | 6 | US-10-225-066A-359 | Sequence 359, App |
| 9 | 286 | 43.6 | 578 | 7 | US-10-374-780A-2719 | Sequence 2719, Ap |
| 10 | 286 | 43.6 | 578 | 9 | US-10-225-066A-359 | Sequence 359, App |
| 11 | 278 | 42.4 | 732 | 8 | US-10-697-787-18 | Sequence 18, Appl |
| c 12 | 276 | 42.1 | 466 | 7 | US-10-260-238-4786 | Sequence 4786, Ap |
| c 13 | 275.5 | 42.0 | 543 | 7 | US-10-021-323-16167 | Sequence 16167, A |
| 14 | 275.5 | 42.0 | 779 | 8 | US-10-767-795-271 | Sequence 271, App |
| 15 | 270.5 | 41.2 | 507 | 7 | US-10-424-599-5904 | Sequence 5904, Ap |
| 16 | 270.5 | 41.2 | 507 | 7 | US-10-425-114-9021 | Sequence 9021, Ap |
| 17 | 266 | 40.5 | 636 | 8 | US-10-767-795-5934 | Sequence 5934, Ap |
| 18 | 266 | 40.5 | 679 | 7 | US-10-424-599-103390 | Sequence 103390, |
| 19 | 261 | 39.8 | 666 | 7 | US-10-424-599-92268 | Sequence 92268, A |
| c 20 | 259.5 | 39.6 | 505 | 7 | US-10-021-323-16005 | Sequence 16005, A |
| 21 | 259.5 | 39.6 | 532 | 7 | US-10-021-323-14482 | Sequence 14482, A |
| 22 | 259.5 | 39.6 | 538 | 7 | US-10-021-323-11476 | Sequence 11476, A |
| c 23 | 259.5 | 39.6 | 542 | 7 | US-10-021-323-11563 | Sequence 11563, A |
| 24 | 259.5 | 39.6 | 604 | 8 | US-10-767-795-5000 | Sequence 5000, Ap |
| 25 | 254.5 | 38.8 | 723 | 8 | US-10-425-115-165252 | Sequence 165252, |
| 26 | 249 | 38.0 | 577 | 7 | US-10-021-323-13836 | Sequence 13836, A |
| 27 | 247 | 37.7 | 737 | 7 | US-10-767-701-20136 | Sequence 20136, A |
| 28 | 246.5 | 37.6 | 424 | 7 | US-10-767-701-27843 | Sequence 27843, A |
| 29 | 246 | 37.5 | 336 | 7 | US-10-437-963-43786 | Sequence 43786, A |
| 30 | 245 | 37.3 | 513 | 5 | US-10-295-403-63 | Sequence 63, Appl |
| 31 | 245 | 37.3 | 513 | 7 | US-10-412-6998-661 | Sequence 661, App |
| 32 | 240.5 | 36.7 | 613 | 7 | US-10-424-599-50010 | Sequence 50010, A |
| 33 | 239 | 36.4 | 273 | 7 | US-10-437-963-48898 | Sequence 48898, A |
| 34 | 239 | 36.4 | 776 | 7 | US-10-260-238-433 | Sequence 433, App |
| 35 | 233.5 | 35.6 | 980 | 7 | US-10-425-114-26727 | Sequence 26727, A |
| 36 | 233.5 | 35.6 | 1056 | 8 | US-10-425-115-39009 | Sequence 39009, A |
| 37 | 228.5 | 34.8 | 584 | 7 | US-10-021-323-9298 | Sequence 9298, Ap |
| 38 | 228 | 34.6 | 932 | 7 | US-10-424-599-56642 | Sequence 56642, A |
| 39 | 227 | 34.6 | 1062 | 8 | US-10-425-115-32995 | Sequence 32995, A |
| 40 | 225.5 | 34.4 | 717 | 8 | US-10-425-115-5670 | Sequence 5670, Ap |
| 41 | 225.5 | 34.4 | 766 | 8 | US-10-425-115-5671 | Sequence 5671, Ap |
| 42 | 222.5 | 33.9 | 418 | 7 | US-10-767-701-30624 | Sequence 30624, A |
| 43 | 220.5 | 33.6 | 387 | 7 | US-10-767-701-18162 | Sequence 18162, A |
| 44 | 218.5 | 33.3 | 330 | 7 | US-10-437-963-55684 | Sequence 55684, A |
| 45 | 214.5 | 32.7 | 342 | 7 | US-10-437-963-58875 | Sequence 58875, A |

ALIGNMENTS

RESULT 1

US-10-697-787-1
; Sequence 1, Application US/10697787
; Publication No. US20040216182A1
; GENERAL INFORMATION:
; APPLICANT: Agrinomics, LLC
; TITLE OF INVENTION: Generation of Plants with Improved Pathogen Resistance and
; DROUGHT TOLERANCE
; FILE OF INVENTION: Drought Tolerance
; FILE REFERENCE: AG03-071C
; CURRENT APPLICATION NUMBER: US/10/697,787
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/US03/12981
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-697-787-1

| | |
|------------------------|----------|
| Alignment Scores: | |
| Pred. No.: | 1.09e-78 |
| Score: | 656.00 |
| Match: | 126 |
| Percent Similarity: | 100.0% |
| Best Local Similarity: | 100.0% |
| Query Match: | 100.0% |
| Indels: | 0 |
| Gaps: | 0 |

US-10-697-787-2 (1-126) x US-10-697-787-1 (1-381)

```
QY 1 MetAlaSerAenSerArgSerSerIleSerProTrrPThrPheSerGlnAenLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAenValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCAATGTGGCAAAA 120
QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGGAGGGAACCTGTAGAAGAGTGAAGCGCCACCTATGACATTTCTCGTCGAGGAT 180
QY 61 LeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGluSerAen 80
Db 181 CTATCTCAACATCGAGACTGGTGTGCTCTTGGCCCAATTACAAGACCTTCGAATCTAAC 240
QY 81 SerArgSerIleAenAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGACACAAGGTATATAACTAAATATCTATATATGATGCTC 300
QY 101 SerIleTyrPheAenHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATAATCATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTTGTA 360
QY 121 SerTyrIleSerLeuVal 126
Db 361 AGTATATTTCTTTGGTT 378
```

RESULT 2

US-10-512-600-1

; Sequence 1, Application US/10512600

; Publication No. US20050210546A1

; GENERAL INFORMATION:

; APPLICANT: AGRINOMICS, LLC

; TITLE OF INVENTION: GENERATION OF PLANTS WITH IMPROVED PATHOGEN RESISTANCE

; FILE REFERENCE: AG03-033C-US

; CURRENT APPLICATION NUMBER: US/10/512,600

; CURRENT FILING DATE: 2004-10-22

; PRIOR APPLICATION NUMBER: US 60/375,333

; PRIOR FILING DATE: 2002-04-24

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-512-600-1

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1,09e-78 | Length: | 381 |
| Score: | 656.00 | Matches: | 126 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-10-697-787-2 (1-126) x US-10-512-600-1 (1-381)

```
QY 1 MetAlaSerAenSerArgSerSerIleSerProTrrPThrPheSerGlnAenLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTACCATGGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAenValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCAATGTGGCAAAA 120
QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGGAGGGAACCTGTAGAAGAGTGAAGCGCCACCTATGACATTTCTCGTCGAGGAT 180
QY 61 LeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGluSerAen 80
```

```
Db 181 CTCATCAACATCGAGACTGGTGTGCTCTTTGCCCAATTACAAGACCTTCGAATCTAAC 240
QY 81 SerArgSerIleAenAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGACACAAGGTATATAACTAAATATCTATATATGATGCTC 300
QY 101 SerIleTyrPheAenHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATAATCATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTTGTA 360
QY 121 SerTyrIleSerLeuVal 126
Db 361 AGTATATTTCTTTGGTT 378
```

RESULT 3

US-10-424-599-6644

; Sequence 6644, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 6644

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_106008C.1

US-10-424-599-6644

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 3,76e-29 | Length: | 678 |
| Score: | 293.00 | Matches: | 69 |
| Percent Similarity: | 63.6% | Conservative: | 20 |
| Best Local Similarity: | 49.3% | Mismatches: | 33 |
| Query Match: | 44.7% | Indels: | 18 |
| DB: | 7 | Gaps: | 4 |

US-10-697-787-2 (1-126) x US-10-424-599-6644 (1-678)

```
QY 1 MetAlaSerAenSer-----ArgSerSerIleSerProTrrPThrPheSerGln 16
Db 206 ATGGCCTCAAGCTCTCTCAGCAAAACAAAGGCTTCTGACTCCTCTTGGAGCCCAAAACAG 265
QY 17 AsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHis 36
Db 266 AACAAAGCTGTTGAAAAGCACCTTGCAAAATATGACAAAGGATACCCCTGAGCGTGCAG 325
QY 37 AsnValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIle 56
Db 326 AATGTAGCCAAAGCAGTAGGTGGAAAATCTGCAGATGAAGTTAAGACACACTATGAAATC 385
QY 57 LeuValGluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThr 76
Db 386 CTCTTGGAGGATCTCAGACACATAGAGTCTGGCGCGTTCCTCTTCCCAAGTACAAAG--- 442
QY 77 PheGluSerAenSerArgSerIleAenAspPheAspThrArgTyrIleThrLysTyrLeu 96
Db 443 -----TCCACAGAGACAGCACCAATGTTGATGAAGACAGAGGCTTCTGAAGTACTT 496
QY 97 -----TyrMetMetLeuSerIleTyrPheAspAsn 106
Db 497 AAACCTGAATTGATGAAGCAACGCTTTATGTTATCTCTGAGCTATTGGCAGCATTTTATG 556
QY 107 HisSerSerAspPheGluLysPheSerGlnLysValLeu---ValSerTyrIleSerLeu 125
```


QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluApp 60
DB 267 GCTGTTGGTGGCAAACTCCAGAGGAAGTGAAAGGCACTACGAACCTCTGTGTCAGGAT 326
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80
DB 327 GTTAAGCATATTGAGTCTGGACGCTGTGCCATTCCCAAAATTACAAGAAAACTACTTCAGGG 386
QY 81 Ser 81
DB 387 TCA 389

RESULT 7

US-09-770-152-473/c
; Sequence 473, Application US/09770152
; Publication No. US20020040489A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2025US (PARA-014PRV)
; CURRENT APPLICATION NUMBER: US/09/770,152
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,503
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 473
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-152-473

Alignment Scores:
Pred. No.: 2,46e-28 Length: 556
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x US-09-770-152-473 (1-556)

QY 1 MetAlaSerAsnSerArgSerSeriIle-----SerProTrrPThrPheSerGlnAsnLys 18
DB 474 ATGCATCAGGCTCAAGTCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 415
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 414 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAACAGACACTCCGGACCGTGGCACAATGTT 355
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
DB 354 GCTAGAGCTGTGGTGGTAAACACCAGAGAGCTAAGACAGTATGACCTTCTAGTT 295

QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
DB 294 CGTGACATCGAAAGCATCGAGAATGGTCACGTGCCATTCCCTGACTACAAGACTACTACA 235
QY 79 SerAsnSer 81
DB 234 GGAAACAGC 226

RESULT 8

US-10-225-066A-359
; Sequence 359, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omlair
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 359
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-359

Alignment Scores:
Pred. No.: 2,61e-28 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 6 Gaps: 1

US-10-697-787-2 (1-126) x US-10-225-066A-359 (1-578)

QY 1 MetAlaSerAsnSerArgSerSeriIle-----SerProTrrPThrPheSerGlnAsnLys 18
DB 108 ATGCATCAGGCTCAAGTCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 167
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 168 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAACAGACACTCCGGACCGTGGCACAATGTT 227
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
DB 228 GCTAGAGCTGTGGTGGTAAACACCAGAGAGCTAAGACAGTATGACCTTCTAGTT 287
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAAGCATCGAGAAATGGTCAGTGCCTCCATTCCTGACTACAAAGACTACTACA 347

QY 79 SerAsnSer 81

Db 348 GGAAACAGC 356

RESULT 9

US-10-374-780A-2719

; Sequence 2719, Application US/10374780A

; Publication No. US20040019927A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James

; APPLICANT: Broun, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10/374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR APPLICATION NUMBER: 09/837,944

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 10/225,066

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,067

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,068

; PRIOR FILING DATE: 2002-08-09

; NUMBER OF SEQ ID NOS: 2906

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2719

; LENGTH: 578

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G1789

US-10-374-780A-2719

Alignment Scores:
Pred. No.: 2,61e-28 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 7 Gaps: 1

US-10-697-787-2 (1-126) x US-10-374-780A-2719 (1-578)

QY 1 MetAlaSerAsnSerArgSerSerIle-----SerProTrrPThrPheSerClnAsnLys 18

Db 108 ATGGCATCAGGCTCAATGCTCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGAACAA 167

QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHisAsnVal 38

Db 168 GCCTTTGAGCGTCTAGCAGTCTATGACCAAGACACTCCGGACCGTTGGCAACAATGTT 227

QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58

Db 228 GCTAGAGCTGTTGGTGAACACACAGAAAGAGCTTAAGAGACAGTATGACCTTCTAGTT 287

QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAAGCATCGAAGTGGTCAGTGCCATTCCTGACTACAAAGACTACTACA 347

QY 79 SerAsnSer 81

Db 348 GGAAACAGC 356

RESULT 10

US-10-225-066A-359

; Sequence 359, Application US/10225066A

; Publication No. US20050160493A9

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omaira

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROWN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MBI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 359

; LENGTH: 578

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-225-066A-359

Alignment Scores:
Pred. No.: 2,61e-28 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 9 Gaps: 1

US-10-697-787-2 (1-126) x US-10-225-066A-359 (1-578)

QY 1 MetAlaSerAsnSerArgSerSerIle-----SerProTrrPThrPheSerClnAsnLys 18

Db 108 ATGGCATCAGGCTCAATGCTCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGAACAA 167

QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHisAsnVal 38

Db 168 GCCTTTGAGCGTCTAGCAGTCTATGACCAAGACACTCCGGACCGTTGGCAACAATGTT 227

```
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspLleLeuVal 58
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 228 GCTAGAGCTGTTGGTGGTAAACACCAAGAGAGCTAAGAGACAGTATGACCTCTTAGTT 287

QY 59 GluAspLeuLleAsnLleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 288 CGTGACATCGAAAGCATCGAAGATGGTCAGTGCCCATTCCTGACTACAAGACTACTACA 347

QY 79 SerAsnSer 81
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 348 GGAACAGC 356

RESULT 11
US-10-697-787-18
; Sequence 18, Application US/10697787
; Publication No. US20040216182A1
; GENERAL INFORMATION:
; APPLICANT: Agrinomics, LLC
; TITLE OF INVENTION: Generation of Plants with Improved Pathogen Resistance and
; FILE OF INVENTION: Drought Tolerance
; CURRENT APPLICATION NUMBER: US/10/697,787
; PRIOR FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/US03/12981
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-10-697-787-18

Alignment Scores:
Pred. No.: 4.54e-27 Length: 732
Score: 278.00 Matches: 69
Percent Similarity: 54.7% Conservative: 19
Best Local Similarity: 42.9% Mismatches: 35
Query Match: 42.4% Indels: 38
DB: 8 Gaps: 6

US-10-697-787-2 (1-126) x US-10-697-787-18 (1-732)

QY 1 MetAlaSerAsnSerArgSerIleSerProTyrPheSerGlnAsnLysMetPhe 20
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 ATGGCATCAAGCTCACTTCATCT--TCATCATGGACACCGCAGCAACAACTATT 61

QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisAsnValAlaLys 40
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 GAAAGGCGACTAGCTCAATCGATAAGGACACACCTGACCGGTGGCAGAAATGTGGCGCG 121

QY 41 AlaVal---GlyGlyLysThrValGluGluValLysArgHisTyrAspLleLeuValGlu 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 CGCGTTTGGAGGTGGAAAATCCGCGATGAAGTAAAGAGACACTATGAATACTATTAG 181

QY 60 AspLeuLleAsnLleGluThrGlyArgValProLeuProAsnTyr----- 74
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 GATCTCAGGCGCATTTGAATCTGGACGCTTCCTTCTTACTTACCCCATGAACACAA 241

QY 74 ----- 74

Db 242 AGGTATTCTTAATCATCTCTTTAAAGTCTTTTGTCCGTATTATTTAAATAATCAACATT 301

QY 75 ---LysThrPheGluSerAsnSerArgSerIleAsnAspPheAspThr-----Arg 90
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 CAAAAGTCTCTTCAAAATTCATTCGATGGAGTGAATAAATATGATATTTTGTTCAAA 361

QY 91 TyrIleThrLysTyrLeuTyrMetMetLeuSerIleTyrPheAspAsnHisSerAsp 110
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 GGAATAGCAAGATATATATCTTTGAATCTTGAA---CATTTTGAATCTGAATGAGACG 418
```

```
QY 111 PheGluLysPheSerGlnLysValLeuVal-----SerTyrIleSer 124
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 GTTCCATACTTAAACCCCTACTTTACTAGTCTATATCTTTGAATGAGACAGTATCATATTT 478

QY 125 Leu 125
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 CTA 481

RESULT 12
US-10-260-238-4786/c
; Sequence 4786, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4786
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Musa acuminata
US-10-260-238-4786

Alignment Scores:
Pred. No.: 4.26e-27 Length: 466
Score: 276.00 Matches: 66
Percent Similarity: 60.0% Conservative: 18
Best Local Similarity: 47.1% Mismatches: 27
Query Match: 42.1% Indels: 30
DB: 7 Gaps: 4

US-10-697-787-2 (1-126) x US-10-260-238-4786 (1-466)

QY 1 MetAlaSerAsnSerArgSerIleSerProTyrPheSerGlnAsnLysMetPhe 20
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 425 ATGGCAACCACTCGGGAGCTCGTGC-----TGGACCGCGAAGCAGAACAAAGATGTC 372

QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisAsnValAlaLys 40
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 371 GAGAAAAGCCCTCCGCGGTGTACGACAAAGGACACCGCCGACCGCTGGCAACACGTGGCGGC 312

QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspLleLeuValGluAsp 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 GCGCTGGCGGCAAGTCAGCGGGAAGAGTGAAGCGGCACCTACGAGCTGCTGGTGGCGGAC 252

QY 61 LeuLleAsnLleGluThrGlyArgValProLeuProAsnTyr----- 74
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 ATTATCTCATCGAAGGCGCAGATGCTCGAGCCAAATTACCGCTCTCCGGCCACAGG 192

QY 75 -----LysThrPheGluSerAsnSerArgSerIleAsnAspPhe 87
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 GGATGAAGTCTCGACTGGATCAAAAGTGTAGAACCGGGCAGCAG-TCTACTCTCGATGAT 133

QY 88 AspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIleTyrPheAspAsnHis 107
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 132 CAATTACAAGAAACAGATACATACATATATGTGATG----- 97
QY 108 SerSerAspPheGluLysPheSerGlnLysValLeu---ValSerTyrIleSerLeuVal 126
Db 96 -----AAGTATCAGACATCCCTGCTTAAGTAGTGTCTGTTCTCTGTC 52

RESULT 13
US-10-021-323-16167/c
; Sequence 16167, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 16167
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-015-Q1-N6-E2
US-10-021-323-16167

Alignment Scores:
Pred. No.: 6,28e-27 Length: 543
Score: 275.50 Matches: 54
Percent Similarity: 80.0% Conservative: 10
Best Local Similarity: 67.5% Mismatches: 13
Query Match: 42.0% Indels: 3
DB: 7 Gaps: 1

US-10-697-787-2 (1-126) x US-10-021-323-16167 (1-543)

QY 7 SerSerIleSer-----ProTrrPheSerGlnAsnLysMetPheGluArgAla 23
Db 527 TCATCGATTTCATATGCTGTTTCATGCAGCAGCAAGCAAAACAAAGATTTCGAAAGGGCT 468
QY 24 LeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaValGly 43
Db 467 TTAGCTGTTTACGACAAGGACACACAGATCGTTGGTACAAATGTTGCTAAAGCTGTGGGA 408
QY 44 GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
Db 407 GAGAAACTGTTGAGGAGTGAAGAGCACTATGAGCTTCTTTGAGATGTTAGACGC 348
QY 64 IleGluThrGlyArgValProLeuProAspTyrLysThrPheGluSerAsnSerArgSer 83
Db 347 ATCGAGTCGGGTGCGGTTCTTTCCCGACTATTGCGACCGTTACCGGGAATCGTCAAGCA 288

RESULT 14
US-10-767-795-271
; Sequence 271, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 271
; LENGTH: 779
```

```
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C42883_1
US-10-767-795-271

Alignment Scores:
Pred. No.: 1,09e-26 Length: 779
Score: 275.50 Matches: 54
Percent Similarity: 80.0% Conservative: 10
Best Local Similarity: 67.5% Mismatches: 13
Query Match: 42.0% Indels: 3
DB: 8 Gaps: 1

US-10-697-787-2 (1-126) x US-10-767-795-271 (1-779)

QY 7 SerSerIleSer-----ProTrrPheSerGlnAsnLysMetPheGluArgAla 23
Db 150 TCATCGATTTCATATGCTGTTTCATGCAGCAGCAAGCAAAACAAAGATTTCGAAAGGGCT 209
QY 24 LeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaValGly 43
Db 210 TTAGCTGTTTACGACAAGGACACACAGATCGTTGGTACAAATGTTGCTAAAGCTGTGGGA 269
QY 44 GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
Db 270 GAGAAACTGTTGAGGAGTGAAGAGCACTATGAGCTTCTTTGAGATGTTAGACGC 329
QY 64 IleGluThrGlyArgValProLeuProAspTyrLysThrPheGluSerAsnSerArgSer 83
Db 330 ATCGAGTCGGGTGCGGTTCTTTCCCGACTATTGCGACCGTTACCGGGAATCGTCAAGCA 389

RESULT 15
US-10-424-599-5904
; Sequence 5904, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 5904
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10533C.1
US-10-424-599-5904

Alignment Scores:
Pred. No.: 2,69e-26 Length: 507
Score: 270.50 Matches: 55
Percent Similarity: 66.7% Conservative: 13
Best Local Similarity: 53.9% Mismatches: 31
Query Match: 41.2% Indels: 3
DB: 7 Gaps: 1

US-10-697-787-2 (1-126) x US-10-424-599-5904 (1-507)

QY 3 SerAsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPheGluArg 22
Db 42 GCTACAACAATGGCTCAAGCCAGCGGTGGACTCCGAGGACAGACAAGAGATTTCAGAAAT 101
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaVal 42
Db 102 GCCCTTGCCATCTTCGACAAGGACACCCACAGATAGTGGCAGACCGTGGCCAGGCGCGTC 161
```

```

Qy      43  GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db      162  GGAGGAAAAACGGTGGAGGAAGTCAAAAGGCATTATGAGAAAGCTCGTGGAAAGATGTCAAG 221

Qy      63  AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg 82
Db      222  AAGATAGAGGAAGGTCAACGTGCCCTCCCCCAATTACCGA-----AGTGCTGCAAGA 272

Qy      83  SerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIle 102
Db      273  GGCTACGGTTACATGGATGAAGAAACCAAGGATGAAGGCTCTGAGCTGCGCAGTGAAGTGTG 332

Qy      103  TyrPhe 104
Db      333  ATTTAC 338

```

Search completed: March 10, 2006, 19:49:25
Job time : 659 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:26:25 ; Search time 151 Seconds
(without alignments)
1483.263 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNRRSSISPTWFSQNMKFM.....HSSDFEKSQKLVSYISLV 126

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl
-Q/abs/ABSSWEB_spool/US10697787/runat_09032006_150631_3419/app_query.fasta_1
-DB=Issued Patents NA -QFT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abs05p -USER=US10697787 @CGN 1.1.290 @runat_09032006_150631_3419
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 195 | 29.7 | 357 | 3 | US-09-640-211A-1581 |
| 2 | 193 | 29.4 | 574 | 3 | US-09-640-211A-1534 |
| 3 | 166 | 25.3 | 383 | 3 | US-09-640-211A-1310 |
| 4 | 131.5 | 20.0 | 532 | 3 | US-09-640-211A-1284 |
| 5 | 124 | 18.9 | 1860 | 3 | US-09-919-497-50 |
| 6 | 124 | 18.9 | 2069 | 3 | US-09-949-016-5801 |
| 7 | 118 | 18.0 | 19112 | 3 | US-09-949-016-17541 |
| 8 | 118 | 18.0 | 36546 | 3 | US-09-949-016-17543 |
| 9 | 117.5 | 17.9 | 345 | 3 | US-09-640-211A-1308 |

| | | | | | | |
|----|------|------|---------|---|---------------------|--------------------|
| 10 | 91 | 13.9 | 332 | 3 | US-09-640-211A-1824 | Sequence 1824, Ap |
| 11 | 85 | 13.0 | 494 | 3 | US-09-270-767-31207 | Sequence 31207, A |
| 12 | 85 | 13.0 | 539 | 3 | US-09-270-767-14937 | Sequence 14937, A |
| 13 | 84.5 | 12.9 | 2274 | 3 | US-09-270-767-10919 | Sequence 10919, A |
| 14 | 81 | 12.3 | 450 | 3 | US-09-640-211A-2026 | Sequence 2026, Ap |
| 15 | 81 | 12.3 | 541 | 3 | US-09-640-211A-255 | Sequence 255, App |
| 16 | 80 | 12.2 | 36181 | 3 | US-08-311-731A-120 | Sequence 120, App |
| 17 | 79.5 | 12.1 | 1830121 | 3 | US-09-557-884-1 | Sequence 1, Appli |
| 18 | 79.5 | 12.1 | 1830121 | 3 | US-09-557-884-1 | Sequence 1, Appli |
| 19 | 79.5 | 12.1 | 1830121 | 3 | US-09-557-884-1 | Sequence 1, Appli |
| 20 | 79 | 12.0 | 1612 | 3 | US-10-158-865-1 | Sequence 1, Appli |
| 21 | 77.5 | 11.8 | 1971 | 3 | US-09-640-211A-2104 | Sequence 2104, Ap |
| 22 | 77 | 11.7 | 1086 | 3 | US-09-248-796A-1108 | Sequence 1108, Ap |
| 23 | 76.5 | 11.7 | 373 | 3 | US-09-134-000C-2502 | Sequence 2502, Ap |
| 24 | 76.5 | 11.7 | 3759 | 2 | US-09-640-211A-1486 | Sequence 1486, Ap |
| 25 | 76.5 | 11.7 | 3759 | 2 | US-08-542-921-1 | Sequence 1, Appli |
| 26 | 76.5 | 11.7 | 3759 | 2 | US-08-880-685-1 | Sequence 1, Appli |
| 27 | 76 | 11.6 | 2502 | 3 | US-08-880-684-1 | Sequence 1, Appli |
| 28 | 74 | 11.3 | 562 | 3 | US-09-543-681A-712 | Sequence 712, App |
| 29 | 74 | 11.3 | 919 | 3 | US-09-640-211A-1552 | Sequence 1552, Ap |
| 30 | 73.5 | 11.2 | 1664976 | 3 | US-09-640-211A-485 | Sequence 485, App |
| 31 | 73.5 | 11.2 | 1664976 | 3 | US-08-916-421B-1 | Sequence 1, Appli |
| 32 | 72.5 | 11.1 | 921 | 3 | US-09-692-570-1 | Sequence 1, Appli |
| 33 | 72.5 | 11.1 | 1815 | 3 | US-09-134-000C-3196 | Sequence 3196, Ap |
| 34 | 72 | 11.0 | 29350 | 3 | US-09-614-221A-213 | Sequence 213, App |
| 35 | 72 | 11.0 | 29350 | 3 | US-09-949-016-11963 | Sequence 11963, A |
| 36 | 71.5 | 10.9 | 1710 | 3 | US-09-949-016-17160 | Sequence 17160, A |
| 37 | 71.5 | 10.9 | 2160 | 2 | US-09-350-729A-4 | Sequence 4, Appli |
| 38 | 71.5 | 10.9 | 2160 | 6 | US-08-082-849B-30 | Sequence 30, Appli |
| 39 | 71.5 | 10.9 | 2211 | 3 | PCT-US94-01624-30 | Sequence 30, Appli |
| 40 | 71.5 | 10.9 | 2292 | 3 | US-09-350-729A-2 | Sequence 2, Appli |
| 41 | 71.5 | 10.9 | 2295 | 3 | US-09-350-729A-3 | Sequence 3, Appli |
| 42 | 71.5 | 10.9 | 2709 | 2 | US-09-350-729A-1 | Sequence 1, Appli |
| 43 | 71.5 | 10.9 | 2709 | 2 | US-08-021-601-11 | Sequence 11, Appli |
| 44 | 71.5 | 10.9 | 2709 | 6 | US-08-082-849B-11 | Sequence 11, Appli |
| 45 | 71.5 | 10.9 | 4235 | 2 | PCT-US94-01624-11 | Sequence 11, Appli |
| | | | | | US-08-021-601-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-09-640-211A-1581
; Sequence 1581, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1581

Alignment Scores:
Pred. No.: 2,548-17
Score: 195.00
Percent Similarity: 67.1%
Best Local Similarity: 54.4%
Query Match: 29.7%
DB: 3
Length: 357
Matches: 43
Conservative: 10
Mismatches: 22
Indels: 5
Gaps: 1

US-10-697-787-2 (1-126) x US-09-640-211A-1581 (1-357)

Qy 3 SerAanSerArgSerIleSerProtrpThrPheSerGlnAsnLysMetPheGluArg 22

Db 76 TCAAGTTTGGATTCATCTGGTTGACTATACACAGAACAAATTTTGAAT 135
Qy 23 AlalaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaVal 42
Db 136 GCTCTAGCTGATTTGATAAGACACCCAGATAAATGGAGAAAGTGGCAGCGGCTG 195
Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspLeuValGluAspLeu 62
Db 196 CCTGGAATAACTGCTACGGATGTTAGAAAGCATTTATGAAGATCTCGTGGAGATGTTACT 255
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSer 81
Db 256 TGTATTGAAGTGC-CGCGTTGCCCTACCCACGTAC-----AGTAACTCT 299

RESULT 2

US-09-640-211A-1534/c
; Sequence 1534, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: Modification of Gene Transcription
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1534
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1534

Alignment Scores:
Pred. No.: 9.31e-17 Length: 574
Score: 193.00 Matches: 39
Percent Similarity: 66.7% Conservative: 13
Best Local Similarity: 50.0% Mismatches: 26
Query Match: 29.4% Indels: 0
DB: 3 Gaps: 0

US-10-697-787-2 (1-126) x US-09-640-211A-1534 (1-574)

Qy 8 SerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyr 27
Db 437 AGTTGTTCTCGTGGAGTAAGAAACAGACAGGCGGTTCGAGAAATGCCCTGGCTACCCAC 378
Qy 28 AspLysAspThrProAspArgTrpHisAsnValAlaValGlyGlyLysThrVal 47
Db 377 TTCGAGGTTTCCCGGATGTTGGAGAGTGGTCCGACGTCGCGGCAAAACCTTA 318
Qy 48 GluGluValLysArgHisTyrAspIleuValGluAspLeuIleGluThrGly 67
Db 317 GAGGAGATTAACTCCACTACGAGGACCTGGTGGAGATGTCAATAGATAGAGGCTGGC 258
Qy 68 ArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArgSerIleAsn 85
Db 257 CGCGTGGCTCTGCCGCTCTACAGTTCGCTTGGCTCGGATGCTCCGCCAGTAAC 204

RESULT 3

US-09-640-211A-1310
; Sequence 1310, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1310
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1310

Alignment Scores:
Pred. No.: 2.4e-13 Length: 383
Score: 166.00 Matches: 29
Percent Similarity: 74.5% Conservative: 9
Best Local Similarity: 56.9% Mismatches: 13
Query Match: 25.3% Indels: 0
DB: 3 Gaps: 0

US-10-697-787-2 (1-126) x US-09-640-211A-1310 (1-383)

Qy 4 AsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAla 23
Db 229 AACGAGGAGGAGGACCGGAAATGACCCCTGCCGAGAACAAATGTTTGGAAAGCG 288
Qy 24 LeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaValGly 43
Db 289 CTGGCGTGTCACGATCAGGACACCGCGATCGGTGGAGTAGGCTCGCGCTCGATGATCCCT 348
Qy 44 GlyLysThrValGluGluValLysArgHisTyr 54
Db 349 GGAAGACGGTGGAGGATGTGTTAAGCACTAT 381

RESULT 4

US-09-640-211A-1284
; Sequence 1284, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: Modification of Gene Transcription
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1284
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1284

Alignment Scores:
Pred. No.: 1.82e-08 Length: 532
Score: 131.50 Matches: 28
Percent Similarity: 57.0% Conservative: 17
Best Local Similarity: 35.4% Mismatches: 33
Query Match: 20.0% Indels: 1
DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x US-09-640-211A-1284 (1-532)

Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 171 TGGAGCGGTACAGGACAAAGGCTTCGACACGCGCTGTGGCGGTGGCGGAGACTCG 230
Qy 32 ProAspArgTrpHisAsnValAlaValGlyLysThrValGluGluValLys 51
Db 231 CCCGACCGGTGGCAGCTGATCGGAAACCGGCTG---AACCGGTCCGGTCGCAAGTCTC 287

QY 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
Db 288 GAGCACTACCAGAGCTGGTGGAGGACATTGACCGATCGAGTCGGGGGGTTCGAGCCG 347
QY 72 ProAsnTyrLysThrPheGluSerAsnSerArgSerIleAsnAspPheAspThrArg 90
Db 348 CCGAGTACCGGACGACACCCCGGCGAGCTGGCGGCGAGATTGCGCTTCGAGACGAAG 404

RESULT 5

US-09-919-497-50
; Sequence 50, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-50

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.12e-06 | Length: | 1860 |
| Score: | 124.00 | Matches: | 21 |
| Percent Similarity: | 64.2% | Conservative: | 13 |
| Best Local Similarity: | 39.6% | Mismatches: | 19 |
| Query Match: | 18.9% | Indels: | 0 |
| DB: | 3 | Gaps: | 0 |

US-10-697-787-2 (1-126) x US-09-919-497-50 (1-1860)

QY 7 SerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaVal 26
Db 1530 ACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGACAGCTTTGAAACAA 1589
QY 27 TyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaValGlyLysThr 46
Db 1590 TACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAGCGGTGCTGCGAGGACA 1649
QY 47 ValGluGluValLysArgHisTyrAspIleLeuValGlu 59
Db 1650 AAGAAGGACTGCATGAACACGATACAAGGAACCTTGTGCGAG 1688

RESULT 6

US-09-949-016-5801
; Sequence 5801, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5801
; LENGTH: 2069
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-5801

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 1.3e-06 | Length: | 2069 |
| Score: | 124.00 | Matches: | 21 |
| Percent Similarity: | 64.2% | Conservative: | 13 |
| Best Local Similarity: | 39.6% | Mismatches: | 19 |
| Query Match: | 18.9% | Indels: | 0 |
| DB: | 3 | Gaps: | 0 |

US-10-697-787-2 (1-126) x US-09-949-016-5801 (1-2069)

QY 7 SerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaVal 26
Db 1588 ACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGACAGCTTTGAAACAA 1627
QY 27 TyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaValGlyLysThr 46
Db 1628 TACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAGCGGTGCTGCGAGGACA 1687
QY 47 ValGluGluValLysArgHisTyrAspIleLeuValGlu 59
Db 1688 AAGAAGGACTGCATGAACACGATACAAGGAACCTTGTGCGAG 1726

RESULT 7

US-09-949-016-17541/c
; Sequence 17541, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17541
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17541

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-------|
| Pred. No.: | 0.0002 | Length: | 19112 |
| Score: | 118.00 | Matches: | 35 |
| Percent Similarity: | 41.7% | Conservative: | 28 |
| Best Local Similarity: | 23.2% | Mismatches: | 59 |
| Query Match: | 18.0% | Indels: | 30 |
| DB: | 3 | Gaps: | 2 |

US-10-697-787-2 (1-126) x US-09-949-016-17541 (1-19112)

QY 3 SerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 17658 TCAGGTCCATATACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGACAA 17599
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaVal 42
Db 17598 GCTTTGAAAAACATACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAGCGGTG 17539
QY 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 17538 CTTGGCAGGACAAAGAGGACTGCATGAACAGGAGTTGACAGCTTAAGCGCTTAT 17479
QY 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe----- 77

```
Db 17478 AGATTGTTGATATTTAAATTTGAGCTTGGGTGATATAGCTCATATTTATGTCTATTAAG 17419
Qy 78 -----|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17418 TGGTTAATAATATAGTCTTATTTCAAAACCACTAGTATGGGTCAATATTTATGAGAA 17359
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMet 98
Db 17358 GAAATGAGCAACCTTAGAAACTTTCTAGGCCCTGG-CAGATAATGAAATCAGTATATTTA 17300
Qy 99 MetLeuSer-----|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17299 GTACTATCTATCTAGTGTGCTGAATCTTTACACCAACACTAGATTTAAATGTCTGTT 17240
Qy 114 PheSerGlnLysValLeuValSerTyrIleSer 124
Db 17239 TTTAAATCGCAACTTCTTAAATCACACCTTTCT 17207
RESULT 8
US-09-949-016-17543
; Sequence 17543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17543
; LENGTH: 36546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17543
Alignment Scores:
Pred. No.: 0.000503 Length: 36546
Score: 118.00 Matches: 35
Percent Similarity: 41.7% Conservative: 28
Best Local Similarity: 23.2% Mismatches: 59
Query Match: 18.0% Indels: 30
DB: 3 Gaps: 2
US-10-697-787-2 (1-126) x US-09-949-016-17543 (1-36546)
Qy 3 SerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 33729 TCAGGTCATATACAGACTTCACCCCTTGACACACAGAAACAGAAAGACTTTTGGACAA 33788
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
Db 33789 GCITTTGAAAACATACCCAGTAATAACACTGAAAGATGGGAAAAAATAGCAGAAAGCGTG 33848
Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 33849 CCTGGCAGGACAAGAAGAGACTGTCATCAACGATACACAGGTGTGCACACTTAAGCGTTAT 33908
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe-----77
Db 33909 AGATTGTTGATTTTAAATTTAGCTTGGGTGATATAGCTCATATTTATGTCTATTAAG 33968
Qy 78 -----|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33969 TGGTTAATAATATAGTCTTATTTCAAAACCACTAGTATGGGTCAATATTTATGAGAA 34028
```

```
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMet 98
Db 34029 GAAATGAGCAACCTTAGAAACTTTCTAGGCCCTGG-CAGATAATGAAATCAGTATATTTA 34087
Qy 99 MetLeuSer-----|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34088 GTACTATCTATCTAGTGTGCTGAATCTTTACACCAACACTAGATTTAAATGTCTGTT 34147
Qy 114 PheSerGlnLysValLeuValSerTyrIleSer 124
Db 34148 TTTAAATCGCAACTTCTTAAATCACACCTTTCT 34180
RESULT 9
US-09-640-211A-1308
; Sequence 1308, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1308
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1308
Alignment Scores:
Pred. No.: 7.77e-07 Length: 345
Score: 117.50 Matches: 23
Percent Similarity: 62.1% Conservative: 13
Best Local Similarity: 39.7% Mismatches: 21
Query Match: 17.9% Indels: 1
DB: 3 Gaps: 1
US-10-697-787-2 (1-126) x US-09-640-211A-1308 (1-345)
Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 171 TGGAGCCGCTACGAGACAAAGGTCTTCGAGCACGCGCTGTGGCGGAGGACTCG 230
Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
Db 231 CCCGACCCGTGGCAGCTGATCGGGAACCGCGCTG---AACCGTCCGCGTCCGAAGTGTTC 287
Qy 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgVal 69
Db 288 GAGCACTACCAAGGCTGTGGAGGACATGTACCGCATTCGAGTCGGGGCGGGTC 341
RESULT 10
US-09-640-211A-1824
; Sequence 1824, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 1824
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1824

Alignment Scores:
Pred. No.: 0.00288 Length: 332
Score: 91.00 Matches: 17
Percent Similarity: 84.6% Conservatives: 5
Best Local Similarity: 65.4% Mismatches: 4
Query Match: 13.9% Indels: 0
DB: 3 Gaps: 0

US-10-697-787-2 (1-126) x US-09-640-211A-1824 (1-332)
QY 49 GluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArg 69
Db 4 GAGTGGAGGAGGATTACGAGCTTCTTGTGAGGATGTGACTGTGATTGATGTCGCCG 63
QY 69 ValProLeuProAsnTyr 74
Db 64 GTTGCTTTCCTGCTTAT 81

RESULT 11
US-09-270-767-31207/c
; Sequence 31207, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31207
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31207

Alignment Scores:
Pred. No.: 0.033 Length: 494
Score: 85.00 Matches: 23
Percent Similarity: 45.8% Conservatives: 6
Best Local Similarity: 37.1% Mismatches: 29
Query Match: 13.0% Indels: 4
DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x US-09-270-767-31207 (1-494)
QY 2 AlaSerAsnSerArgSerSerIleSerPro-----TyrThrPheSerGlnAsnLys 18
Db 191 GCTTCAGCGGAGGAGCATGTCTAATTCGGAGACCACTGGAGCAAGCAGCAGCGC 132
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThr---ProAspArgTrpHisAsn 37
Db 131 GCCCTGGAGCGCGCATAGTCAAGTACCGGAGAGACGGCGCGGAGACCGTTGGCAGAG 72
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 71 ATCGCCAAACAGTGTGCGGAGAGACCAAGAGGAGTGCCTGGTGGCTACCAAGTATCTC 12
QY 58 ValGlu 59
Db 11 TCGGAG 6

RESULT 12
US-09-270-767-14937/c
; Sequence 14937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14937
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14937

Alignment Scores:
Pred. No.: 0.0373 Length: 539
Score: 85.00 Matches: 23
Percent Similarity: 46.8% Conservatives: 6
Best Local Similarity: 37.1% Mismatches: 29
Query Match: 13.0% Indels: 4
DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x US-09-270-767-14937 (1-539)
QY 2 AlaSerAsnSerArgSerSerIleSerPro-----TyrThrPheSerGlnAsnLys 18
Db 236 GCTTCAGCGGAGGAGCATGTCTAATTCGGAGACCACTGGAGCAAGCAGCAGCGC 177
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThr---ProAspArgTrpHisAsn 37
Db 176 GCCCTGGAGCGCGCATAGTCAAGTACCGGAGAGACGGCGCGGAGACCGTTGGCAGAG 117
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 116 ATCGCCAAACAGTGTGCGGAGAGACCAAGAGGAGTGCCTGGTGGCTACCAAGTATCTC 57
QY 58 ValGlu 59
Db 56 TCGGAG 51

RESULT 13
US-09-270-767-10919
; Sequence 10919, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10919
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-10919

Alignment Scores:
Pred. No.: 0.338 Length: 2274
Score: 84.50 Matches: 17
Percent Similarity: 53.8% Conservatives: 11
Best Local Similarity: 32.7% Mismatches: 23
Query Match: 12.9% Indels: 1
DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x US-09-270-767-10919 (1-2274)
QY 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 1924 TGGACCAAGGAGGAGCAAGCCCTGCTCGAAGAGCGCATTAATAAACCTATCAACGACGACA 1983
```

```
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14937
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14937
```


GenCore version 5.1.1.7
Copyright (c) 1993 - 2006, Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:19:36 ; Search time 3742 Seconds
(without alignments)
1575.407 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNSRSSISPTWFSQKMF.....HSSDFKFSQKLVLSVSLV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q-/abss/ABSSWEB_spool/US10697787/runat_09032006_150629_3371/app_query.fasta_1
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US10697787 @CGN 1.1 8010 @runat_09032006_150629_3371 -NCPU=6 -ICPU=3
-NO WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 311 | 47.4 | 446 | 8 | L38243 BNAF0581E M |
| 2 | 308 | 47.0 | 728 | 5 | BU890694 P |
| 3 | 305 | 46.5 | 626 | 6 | CB920052 VVD058F11 |
| 4 | 305 | 46.5 | 662 | 6 | CB919185 VVD042D03 |
| 5 | 304 | 46.3 | 474 | 5 | BU668323 MC01026A1 |
| 6 | 299 | 45.6 | 349 | 5 | BU822353 UB36DPB08 |
| 7 | 299 | 45.6 | 624 | 7 | CV232994 WS0199.B2 |

| | | | | | | | |
|---|----|-------|------|-----|---|----------|--------------------|
| C | 8 | 299 | 45.6 | 627 | 7 | CV230387 | CV230387 WS01916.B |
| | 9 | 297 | 45.3 | 256 | 9 | CC179487 | CC179487 SALK_0699 |
| | 10 | 294 | 44.8 | 598 | 1 | AJ768574 | AJ768574 ADJ768574 |
| | 11 | 293.5 | 44.7 | 629 | 8 | DR994824 | DR994824 Mdab9007G |
| | 12 | 292 | 44.5 | 456 | 7 | CO052301 | CO052301 Mdfw20570 |
| C | 13 | 290.5 | 44.3 | 463 | 1 | AV440717 | AV440717 AV440717 |
| C | 14 | 290 | 44.2 | 787 | 7 | CV281695 | CV281695 WS0182.B2 |
| C | 15 | 289 | 44.1 | 906 | 8 | DR552928 | DR552928 WS02616.B |
| | 16 | 288.5 | 44.0 | 407 | 1 | AV442619 | AV442619 AV442619 |
| | 17 | 288 | 43.9 | 497 | 6 | CF603713 | CF603713 BACCA01.0 |
| C | 18 | 288 | 43.9 | 611 | 6 | CF373337 | CF373337 CSECS086D |
| | 19 | 287.5 | 43.8 | 513 | 7 | CO997831 | CO997831 pam01-16m |
| | 20 | 286 | 43.6 | 497 | 1 | AI993653 | AI993653 701496923 |
| C | 21 | 285.5 | 43.5 | 516 | 7 | CV278823 | CV278823 WS0147.B2 |
| C | 22 | 285.5 | 43.5 | 588 | 7 | CV283381 | CV283381 WS0187.B2 |
| | 23 | 285.5 | 43.5 | 662 | 7 | CV277080 | CV277080 WS0142.B2 |
| | 24 | 285 | 43.4 | 457 | 6 | CD669972 | CD669972 eep1c.pk0 |
| | 25 | 283.5 | 43.2 | 599 | 8 | CV882244 | CV882244 Md1v3.403 |
| C | 26 | 283.5 | 43.2 | 653 | 8 | DR997716 | DR997716 Mdfb804M |
| | 27 | 283.5 | 43.2 | 659 | 8 | DR990507 | DR990507 Md1r7005H |
| | 28 | 283.5 | 43.2 | 660 | 8 | DR995864 | DR995864 Mdab9011J |
| C | 29 | 283.5 | 43.2 | 688 | 7 | CV186968 | CV186968 Md1v2.401 |
| | 30 | 283 | 43.1 | 663 | 5 | BU091680 | BU091680 bt76a07.Y |
| C | 31 | 282 | 43.0 | 476 | 6 | CB035850 | CB035850 VVA008A06 |
| | 32 | 282 | 43.0 | 492 | 6 | CB002633 | CB002633 VVB015A08 |
| | 33 | 282 | 43.0 | 492 | 6 | CB002772 | CB002772 VVB020F07 |
| | 34 | 282 | 43.0 | 492 | 6 | CB004008 | CB004008 VVB034G07 |
| | 35 | 282 | 43.0 | 509 | 6 | CB344615 | CB344615 CA48EN000 |
| | 36 | 282 | 43.0 | 513 | 6 | CB342316 | CB342316 CA32EN000 |
| | 37 | 282 | 43.0 | 528 | 6 | CA811699 | CA811699 CA41LN021 |
| | 38 | 282 | 43.0 | 585 | 3 | BM437293 | BM437293 VVA017C08 |
| | 39 | 282 | 43.0 | 618 | 6 | CB344529 | CB344529 CA48EN000 |
| | 40 | 282 | 43.0 | 620 | 6 | CB343867 | CB343867 CA48EN000 |
| | 41 | 282 | 43.0 | 623 | 6 | CB343483 | CB343483 CA32EN000 |
| | 42 | 282 | 43.0 | 624 | 6 | CB341211 | CB341211 CA32EN000 |
| C | 43 | 282 | 43.0 | 636 | 6 | CB343539 | CB343539 CA32EN000 |
| | 44 | 282 | 43.0 | 637 | 6 | CA811769 | CA811769 CA41LN031 |
| | 45 | 282 | 43.0 | 697 | 6 | CB343846 | CB343846 CA32EN000 |

ALIGNMENTS

RESULT 1

L38243
LOCUS L38243
DEFINITION BNAF0581E Mustard flower buds Brassica rapa cdna, mRNA linear EST 03-JUL-1995
ACCESSION L38243
VERSION L38243.1 GI:887283
KEYWORDS EST.
SOURCE Brassica rapa (Brassica campestris)

ORGANISM Brassica rapa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 446)

AUTHORS

Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H.,

Hwang,I. and Cho,M.J.

Expressed sequence tags of Chinese cabbage flower bud cDNA

Plant Physiol. 111 (2), 577-588 (1996)

8787028

COMMENT

Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.

Plant Molecular Biology and Biotechnology Research Center

Email: pmbrc@ongae.gsnu.ac.kr.

FEATURES

Location/Qualifiers

1..446

/organism="Brassica rapa"

/mol_type="mRNA"

/strain="pekinensis"

/db_xref="taxon:3711"

/clone_lib="Mustard flower buds"

/note="Devel_stage = flower bud"

ORIGIN


```

DB:
US-10-697-787-2 (1-126) x CB920052 (1-626)
  1 MetAlaSer-----AnSerArgSerSerIleSerProTTPThrPheSerGlnAsn 17
  148 ATGCATCCAGCTCAATGTCCTCCCGCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAAC 207
  18 LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHiseAsn 37
  208 AAAGCCTTTGAAGAGCGCTTGAGTGTATGACAGGACACCCCTGACCGCTGGTACAT 267
  38 ValAlaLysAlaValGlyGlyThrValGluGluValLysArgHisTyrAspIleLeu 57
  268 GTTGCAGGCGCTGTGGTGGGAAACCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTT 327
  58 ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAspTyrLysThrPhe 77
  328 GTTGAGATATCAAGAGCATCGACTCAGACAAAGTGCCTTCCCCAATTACAAGACTACC 387
  78 GluSerAsnSerArgSer---IleAsnAspPheAspThrArg 90
  388 GGAGCTAGTGGCGGTCCTCAACATGAGCGACAGGAAAGAGG 429

RESULT 4
CB919185          662 bp  mRNA  linear  EST 25-APR-2003
LOCUS
DEFINITION
stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
CDNA clone VVD042D03 5, mRNA sequence.
CB919185          1 GI:30133846
VERSION
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Vitaceae; Vitis.
1 (bases 1 to 662)
Cushman,J.C.
An expressed sequence tag database for abiotic stressed berries of
Vitis vinifera var. Chardonnay
Unpublished (2002)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 042 row: D column: 03
Seq primer: T3 20mer
High quality sequence stop: 662.
FEATURES
    source
        1..662
            /organism="Vitis vinifera"
            /mol_type="mRNA"
            /db_xref="taxon:29760"
            /clone="VVD042D03"
            /tissue_type="berries"
            /dev_stages="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
            /clone_lib="An expressed sequence tag database for abiotic
            stressed berries of Vitis vinifera var. Chardonnay".
            /note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site_1:
            EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.:      2,578-30      Length:      662
Score:          305.00      Matches:      62

```

```

Percent Similarity: 80.9%      Conservative: 14
Best Local Similarity: 66.0%      Mismatches: 14
Query Match: 46.5%      Indels: 4
DB: Gaps: 2

US-10-697-787-2 (1-126) x CB919185 (1-662)

QY 1 MetAlaSer-----AnSerArgSerSerIleSerProTTPThrPheSerGlnAsn 17
Db 148 ATGCATCCAGCTCAATGTCCTCCCGCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAAC 207
QY 18 LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHiseAsn 37
Db 208 AAAGCCTTTGAAGAGCGCTTGAGTGTATGACAGGACACCCCTGACCGCTGGTACAT 267
QY 38 ValAlaLysAlaValGlyGlyThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 268 GTTGCAGGCGCTGTGGTGGGAAACCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTT 327
QY 58 ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAspTyrLysThrPhe 77
Db 328 GTTGAGATATCAAGAGCATCGACTCAGACAAAGTGCCTTCCCCAATTACAAGACTACC 387
QY 78 GluSerAsnSerArgSer---IleAsnAspPheAspThrArg 90
Db 388 GGAGCTAGTGGCGGTCCTCAACATGAGCGACAGGAAAGAGG 429

RESULT 5
BU668323          474 bp  mRNA  linear  EST 31-DEC-2003
LOCUS
DEFINITION
MC01026A10 MC01 Sesamum indicum cDNA, mRNA sequence.
BU668323
ACCESSION
BU668323.1 GI:4046455
VERSION
KEYWORDS
SOURCE
Sesamum indicum (sesame)
ORGANISM
Sesamum indicum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; Lamiales; Pedaliaceae; Sesamum.
1 (bases 1 to 474)
Suh,M.C.
Comparative analysis of Expressed Sequence Tags between Sesamum
indicum and Arabidopsis thaliana developing seeds
Unpublished (2003)
Contact: Suh, Mi Chung
Graduate School of Biotechnology, Korea University
1, 5-Ka, Anam-dong, Sungbuk-ku, Seoul 136-701, Rep. of Korea
Tel: 82 2 3290 3169
Fax: 82 2 927 9028
Email: michung@korea.ac.kr
Plate: 026 row: A column: 10.
FEATURES
    source
        Location/Qualifiers
            1..474
                /organism="Sesamum indicum"
                /mol_type="mRNA"
                /db_xref="taxon:4182"
                /tissue_type="developing seed"
                /cell_line="Yangbaeck"
                /dev_stages="5 to 25 days after flowering"
                /clone_lib="MC01"
                /note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
                XhoI"

ORIGIN
Alignment Scores:
Pred. No.:      2,24e-30      Length:      474
Score:          304.00      Matches:      56
Percent Similarity: 80.2%      Conservative: 13
Best Local Similarity: 65.1%      Mismatches: 13
Query Match: 46.3%      Indels: 4
DB: Gaps: 1

US-10-697-787-2 (1-126) x BU668323 (1-474)

```

```

Qy 1 MetAlaSerAenSer-----ArgSerSerIleSerProTrpThrPheSerGln 16
Db 162 ATGGCGTCGAGTCCATGACATCCGACGTGGCTCAGCCTCAGCGTGACGCTCTAAGCAA 221
Qy 17 AsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHis 36
Db 222 AACAAGCAGTTTCGAGGAGGCTTTGGCAATGTATGACACGACGCGCTGACCGTTGGCAT 281
Qy 37 AsnValAlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIle 56
Db 282 AATATACCCGAGCGCGTGGTGGAGTACGACGAGGAGTGTAGAGGCATTATGAGATT 341
Qy 57 LeuValGluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThr 76
Db 342 CTGTTTAAGACATTATGACATCGACGCGATCAGGTGCGAATACCTAATACTACAGAGCC 401
Qy 77 PheGluSerAenSerArg 82
Db 402 ATGGGAGCAACAGCAGA 419

RESULT 6
LOCUS BU822353
DEFINITION BU822353 349 bp mRNA linear EST 15-OCT-2002
5 prime, mRNA sequence.
ACCESSION BU822353
VERSION BU822353.1 GI:23990180
KEYWORDS EST.
SOURCE Populus tremula
ORGANISM Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 349)
Umeberg,P., Bhalerac,R.R., Jansson,S. and Sterky,F.
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
Unpublished (2002)
JOURNAL Umea Plant Science Center
COMMENT Department of Plant Physiology
University of Umea, 901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: rupali.bhalerao@plantphys.umu.se.

FEATURES
source
1..349
Location/Qualifiers
/organism="Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:113636"
/tissue_type="cambium"
/clone_lib="Populus tremula cambium cDNA library"

ORIGIN
Alignment Scores:
Pred. No.: 7.03e-30 Length: 349
Score: 299.00 Matches: 56
Percent Similarity: 80.5% Conservative: 10
Best Local Similarity: 68.3% Mismatches: 16
Query Match: 45.8% Indels: 0
DB: 5 Gaps: 0

US-10-697-787-2 (1-126) x BU822353 (1-349)

Qy 3 SerAenSerArgSerSerIleSerProTrpThrPheSerGlnAenLysMetPheGluArg 22
Db 99 ACGTCTTCAATGGCTCCGGCTCTCTCTGGACACGCCAAACAAAGCTATTTCGAGAAG 158
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHisAenValAlaLysAlaVal 42
Db 159 GCCCTGGCTGTATACGACAAAGACACCCCGACCGCTGGCAAAATGTGGCAAGCCCGTG 218

```

```

Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 219 GGTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 278
Qy 63 AsnIleGluThrGlyArgValProLeuProAenTyrLysThrPheGluSerAenSerArg 82
Db 279 TACATATAATTCGGCGAAGCGCTCTGCGGAATTACAAAGCCCTCTCGCAGCAATGTTAA 338
Qy 83 SerIle 84
Db 339 GCAGTT 344

RESULT 7
CV232994/c
LOCUS CV232994
DEFINITION CV232994 624 bp mRNA linear EST 21-SEP-2004
WS0199.E24.3', mRNA sequence.
ACCESSION CV232994
VERSION CV232994.1 GI:52389099
KEYWORDS EST.
SOURCE Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM Populus trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 624)
Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
JOURNAL Contact: Joerg Bohmann
COMMENT Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohmann@msl.ubc.ca
Plate: WS0199 row: E column: 24
High quality sequence stop: 624
POLYA=Yes.

FEATURES
source
1..624
Location/Qualifiers
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="VT-125"
/db_xref="taxon:3694"
/clone="WS0199.E24"
/sex="Not determined"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PT-DX-N-A-10"
/notes="Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5'
end of cDNA); Site_2: XhoI (3' end of cDNA); Outer xylem
from 5 year old trees harvested every two weeks between
April and October of 2002 at the University of British
Columbia south campus farm in Vancouver, British Columbia.
mRNA was isolated from each tissue source independently
and equal quantities of mRNA from each tissue were then
pooled. cDNA was prepared from 5 micrograms of mRNA and
directionally ligated into the pBluescript II SK (+) XR
vector using the pBluescript II XR cDNA Library
Construction Kit according to manufacturer's instructions
with modifications (Stratagene) plasmid DNA was then
transformed by electroporation into DH10B cells
(invitrogen) for propagation. Normalization was applied
according to published methods [Bonaldo M.F. et al. (1996)
Genome Research 6(9):791] in order to reduce the abundance
of highly expressed transcripts."

```


COMMENT

Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At1G75250.
 Class: TDNA tagged.

FEATURES

source

Location/Qualifiers

1..256
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Col-0"
 /db_xref="taxon:3702"
 /clone_lib="SALK 069941.39.90.x"
 /clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Alignment Scores:
 Pred. No.: 8.65e-30 Length: 256
 Score: 297.00 Matches: 57
 Percent Similarity: 95.4% Conservative: 5
 Best Local Similarity: 87.7% Mismatches: 3
 Query Match: 45.3% Indels: 0
 DB: 9 Gaps: 0
 US-10-697-787-2 (1-126) x CC179487 (1-256)

Qy 1 MetAlaSerAenSerArgSerSerIleSerProTrpThrPheSerGlnAenLysMetPhe 20
 Db 60 ATGGCGTCAAACTCAAAAGTTCAATCTCACATGGAGCGTTAGTCAAAACAAAGATGTC 119
 Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAenValAlaLys 40
 Db 120 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCATGGCACCATGTGCCAAA 179
 Qy 41 AlaValGlyLysThrValGluGluValLysArgHisTyAspIleLeuValGluAsp 60
 Db 180 GCTGTCGGAGGAAACAGTATAGAGAGTGAACCGCTCTATGACATTCGTCACAAAG 239
 Qy 61 LeuIleAenIleGlu 65
 Db 240 CTCATCAACATCCAG 254

RESULT 10

AJ768574

LOCUS

AJ768574 598 bp mRNA linear EST 06-AUG-2004
 clone P0000100019D06F1, mRNA sequence.

DEFINITION

AJ768574

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Populus euphratica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
 1 (bases 1 to 598)
 Brosche,M., Alacalo,E.R., Vinocur,B., Altman,A., Teichmann,T.,
 Ottow,E.A., Polle,A., Djilianov,D., Afif,D., Triboulot,M.B.,
 Dreyer,E., Paulin,L. and Kangasjaervi,J.
 Gene expression in desert leaves
 Unpublished (2004)
 Contact: Prof. Jaakko Kangasjaarvi

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Department of Biosciences, Plant Physiology

University of Helsinki

FIN-00014 Helsinki, Finland.

Location/Qualifiers

1..598

/organism="Populus euphratica"

/mol_type="mRNA"

/db_xref="taxon:75702"

/clone="P0000100019D06F1"

/tissue_type="leaf"

/dev stage="adult"

/clone_lib="Populus euphratica leaf adult"

/note="country: Israel;Ein Avdat"

ORIGIN

Alignment Scores:

Pred. No.: 6.86e-29 Length: 598
 Score: 294.00 Matches: 56
 Percent Similarity: 80.0% Conservative: 12
 Best Local Similarity: 65.9% Mismatches: 15
 Query Match: 44.8% Indels: 2
 DB: 1 Gaps: 1
 US-10-697-787-2 (1-126) x AJ768574 (1-598)

Qy 1 MetAlaSerAenSerArgSerSerIleSerProTrpThrPheSerGlnAenLys 18

Db 160 ATGGCATCAAGCTCAATGCTCTCGCGTTCGGGCTCGTGGACTGCTCAGAGAACAA 219

Qy 19 MetPheGluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAenVal 38

Db 220 GCCTTTGAAAGGCTTTGGCCGTGTATGACAGAGACACGCCGATCGCTGTACAATGTA 279

Qy 39 AlaLysAlaValGlyLysThrValGluGluValLysArgHisTyAspIleLeuVal 58

Db 280 GCCAGGCGGTGCTGGAAAAACCGCAGAGAGTGAAGAGGCACTATGAACTGCTTGTG 339

Qy 59 GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyLysThrPheGlu 78

Db 340 GAGGATGTGAAGCATATTGAGTGGGTGTCATGTTCTTTCCTTAATTACAGGACTACTGGA 399

Qy 79 SerAenSerArgSer 83

Db 400 GCAATGGCCACGCA 414

RESULT 11

DR994824

LOCUS

DEFINITION

DR994824 629 bp mRNA linear EST 04-AUG-2005

MDae9007G21.g1 Apple EST Mda5 Malus x domestica cDNA similar to

ref|NP_179759.1| myb family transcription factor [Arabidopsis

thaliana] >pir|G84603 hypothetical protein A2921650 [imported] -

Arabidopsis thaliana >gb|AAD23640.1| unknown protein [Arabidopsis

thaliana] >gb|AAP40381.1| putative myb family transcript, mRNA

sequence.

ACCESSION

DR994824

VERSION

KEYWORDS

SOURCE

ORGANISM

Malus x domestica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
 1 (bases 1 to 629)
 Korban,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
 Aldwinckle,H., Malnoy,M., Carroll,N., Goldsborough,P., Orvis,K.,
 Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
 Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
 Tsagareishvili,R., Kennedy,S., Waterston,R. and Wilson,R.
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)
 Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: eat@watson.wustl.edu
Library material provided by S. Korban/H. Aldwinckle/ M. Malnoy
Library constructed by D.O.Gonzalez/J. Vodkin library sequenced by
Washington University Genome Sequencing Center
This trace has been recalled with phred
original value before phred recall for SL was 118
original value before phred recall for SR was 873.

FEATURES

Source

1. 629

/organism="Malus x domestica"

/mol type="mRNA"

/cultivar="GoldRush"

```
/db xref="taxon:3750"
```

```
/tissue type="Leaves challenged with a
```

```

/crsae_type= leaves challenged with a
inaequalia")

```

```

/ lab host="DH10B ampicillin resistant"

```

```
/usr/local/lib/clone lib="Annals EST Md="
```

```

/clone lib="Apple EST_Mdas"
/note="Vector: npTwocount IT cv (1) sites: what

```

```

/note="Vector: pBluescript II SK (+); Site_1: XhoI;
Site_2: EcoRI; Total DNA used out-matched from: 300-350

```

Site_2: EcoRI; Total RNA was extracted from freeze dried

leaf tissue, using the method described by Wang and Vodkin

(Plant Molecular Biology Reporter 12:132-145, 1994).

Poly(A) + mRNA was isolated from total RNA using the

PolyAttract mRNA Isolation system III (Promega). The

library was prepared using the Stratagene pBluescript II

XR cDNA library construction kit. Complementary DNA⁺ was

synthesized from mRNA using a poly (dT) sequence with a

XhoI restriction site. EcoRI adapters were ligated to the

blunt ended cDNA fragments followed by restriction with

XhoI. The cDNA insert is protected from XhoI digestion via

methylation during first strand cDNA synthesis. The cDNA

fragments were directionally cloned into the EcoRI-XhoI

restriction site of the pBluescript vector. The ligated

CDNA fragments were transformed into *E. coli* ElectroMax

DNase I fragments were transformed into E. coli FLECTIONMax DH10B host cells. Transformation efficiency: 1.0E+04

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity

Query Match:

DB:

| | | |
|----------|---------------|-----|
| 8.57e-29 | Length: | 629 |
| 293.50 | Matches: | 62 |
| 63.8% | Conservative: | 12 |
| 53.4% | Mismatches: | 17 |
| 44.7% | Indels: | 25 |
| 8 | Gaps: | 2 |

US-10-697-787-2 (1-126) X DR994824 (1-629)

| | | | | |
|----|-----|--|--|----|
| Qy | 1 | MetAlaSer----- | AsnSerArgSerSerIleSerProTrpThrPheSerGlnAsn | 17 |
| | | | | |
| Db | 87 | ATGGCATCAAGCTCAATGCTCTCAAGAGGCTCTGGCTCTCTCGGACTGCCAAGCAAAAC | 146 | |
| | | | | |
| Qy | 18 | LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsn | 37 | |
| | | | | |
| Db | 147 | AAAGCCCTTTGAAAGGCTCTGGCTGTCTACGACAAGGACACTGCTGACCGCTGGTCAAT | 206 | |
| | | | | |
| Qy | 38 | ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu | 57 | |
| | | | | |
| Db | 207 | GTGGCCAAAGCGGTGGCGGCAAAACCCGGAGGAAGTCAAGAGGCACATGATGAGTGTCTT | 266 | |
| | | | | |
| Qy | 58 | ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe | 77 | |
| | | | | |
| Db | 267 | GTGGAAAGATGTCAGCATATTGAGTCAGGCAGGTGCCCTTCCAGATTACAGGACTACT | 326 | |
| | | | | |
| Qy | 78 | GluSerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyr | 97 | |
| | | | | |
| Db | 327 | GTGGGGAACAGCCATGGC----- | 344 | |
| | | | | |
| Qy | 98 | MetMetLeuSerIleTyrPheAspAsnHisSerSerAspPheGluLys | 113 | |
| | | | | |
| Db | 345 | -----CACCACATGATGATCAGGAAAG | 368 | |

before amplification was 1.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 3 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA remaining partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 9x10⁶ cfu. Background of empty clones was less than 1%."

ORIGIN

Alignment Scores:
Pred. No.: 8,88e-29 Length: 456
Score: 292.00 Matches: 53
Percent Similarity: 87.8% Conservative: 12
Best Local Similarity: 71.6% Mismatches: 9
Query Match: 44.5% Indels: 0
DB: 7 Gaps: 0

US-10-697-787-2 (1-126) x C0052301 (1-456)

QY 3 SerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 60 TCCTTCCTCGAGGAATTCCTCACTCTCTGTCGACCGCCGAGGAGCAACAGCTGTTTCGAGAG 119
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
Db 120 GCCTGCGCGGTGTACGACAGGACACGACGAGCGCTGGCATAATGTGGCCAGGCGCTC 179
QY 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuLe 62
Db 180 GCGCGGAATTCCTGCGGAGGAAGTAAGAGGACACTACGAGATCTCGTGCATGATCTCATG 239
QY 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThr 76
Db 240 CACATAGATCCGCGCATGTCCCTATACCAATTTAATGATCC 281

RESULT 13
AV440717/c
LOCUS AV440717 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION Old Arabidopsis thaliana cDNA clone AP207c09_f 3', mRNA sequence.
ACCESSION AV440717
VERSION AV440717.1 GI:7611088
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 463)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. .463
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

FEATURES
source
1. .463
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

/scotype="Columbia"
/db_xref="taxon:3702"
/clone="AP207c09_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana above-ground organ two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
Pred. No.: 1.45e-28 Length: 463
Score: 290.50 Matches: 59
Percent Similarity: 70.4% Conservative: 22
Best Local Similarity: 51.3% Mismatches: 25
Query Match: 44.3% Indels: 9
DB: 1 Gaps: 2

US-10-697-787-2 (1-126) x AV440717 (1-463)

QY 1 MetAlaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
Db 385 ATGGCTTCCAACCTCAATGAGCTCTAGCGCTTCTTGACACGTAAGGAGAACAAATTTATTT 326
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 325 GAAAGGCGGTGGCTACATATGACAGGACACTCTCGACCGTTGGCATAACGTTGCAAGA 266
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 265 GCGCTTGGCGCAATTCAGCTGAAGAAGTAAGCGGACACTACGAGCTCTCTTAGGGAT 206
QY 61 LeuIleAlaValGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAs 80
Db 205 GTCAATGACATTGAGTCAGGCGGTATTCACACTCCCAATTTACCGTTCAATGGAACAAC 146
QY 80 nSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetL 100
Db 145 CACTGAAAGCATTTAAAGGAACCTCCAA---AAGGCTCTACTGAAGTGG----- 102
QY 100 euSerIleTyrPheAspAsnHisSerSerAspPheGluLys 113
Db 101 -----ACATGGTTCCAAACCATTCCAAAGCACATCTAAG 66

RESULT 14

CV281695 787 bp mRNA linear EST 22-SEP-2004
LOCUS WS0182_B21_J21 PTXD-IL-N-A-9 Populus trichocarpa x Populus
DEFINITION deltoides cDNA clone WS0182_J21 3', mRNA sequence.

ACCESSION CV281695
VERSION CV281695.1 GI:52534670
KEYWORDS EST.
SOURCE Populus trichocarpa x Populus deltoides
ORGANISM Populus trichocarpa x Populus deltoides

REFERENCE 1 (bases 1 to 787)
AUTHORS Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y., Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G., Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Mason, A., Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A., Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K., and Bohlmann, J.

TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@sl.ubc.ca
Plate: WS0182 row: J column: 21
High quality sequence stop: 787
POLYA=Yes.

FEATURES

source

1. 787
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-IL-N-A-9"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma distria Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods (Bonaldo M.F. et al. (1996) Genome Research 6(9):791) in order to reduce the abundance of highly expressed transcripts."

ORIGIN

| Alignment Scores: | |
|------------------------|--|
| Pred. No.: | 3,448-28 |
| Score: | 290.00 |
| Percent Similarity: | 78.8% |
| Best Local Similarity: | 65.9% |
| Query Match: | 44.2% |
| DB: | 7 |
| | US-10-697-787-2 (1-126) x CV281695 (1-787) |
| Length: | 787 |
| Matches: | 56 |
| Conservative: | 11 |
| Mismatches: | 16 |
| Indels: | 2 |
| Gaps: | 1 |

Qy 24 LeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaValGly 43
 Db 617 TTAGCCATTATGACAAAGACACTCTTGACCGTTGGCATAATGTGGCTTCTGTGGTTGGC 558
 Qy 44 GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
 Db 557 GGGAAATCCCGAGGAGGAGTCAAGAGGCATTATGAGATTCTGTGGAGGACTTGAATTCC 498
 Qy 64 IleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArgSer 83
 Db 497 ATTGAAGCAGGCCAGGTGCCTTTCCCAAAATACATATCTTCGAGGCCCTAACACAGGGAG 438
 Qy 84 Ile 84
 Db 437 ATC 435

Search completed: March 10, 2006, 20:29:59
 Job time : 3748 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:14:46 ; Search time 475 Seconds
(without alignments)
1767.897 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASGRSSISPTWFSQNKMF.....HSSDFEKFQKVLVSIVISLV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xl
-Q=/abss/ABSSWEB_spool/US10697787/runat_09032006_150625_3305/app_query.fasta 1
-DB=N Geneseq -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss06p
-USER=US10697787 @CGN 1.1 1096 @runat_09032006_150625_3305 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

N Geneseq 21.1*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 656 | 100.0 | 381 | 10 | ADP43565 | Adf43565 Thale cre |
| 2 | 473 | 72.1 | 294 | 14 | AEA27134 | Aea27134 Stress to |
| 3 | 343 | 52.3 | 279 | 12 | ADO62042 | Ado62042 Transcrip |
| 4 | 327.5 | 49.9 | 478 | 3 | AAC34192 | Aac34192 Arabidops |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 5 | 290 | 44.2 | 760 | 13 | ADR59489 | Adr59489 Cotton cD |
| 6 | 289 | 44.1 | 508 | 3 | AAC35256 | Aac35256 Arabidops |
| 7 | 287 | 43.8 | 704 | 13 | ADX32012 | Adx32012 Plant ful |
| 8 | 286 | 43.6 | 306 | 3 | AAC42715 | Aac42715 Arabidops |
| 9 | 286 | 43.6 | 556 | 10 | ABX57121 | Abx57121 Arabidops |
| 10 | 286 | 43.6 | 578 | 10 | ADD30327 | Add30327 Plant yle |
| 11 | 286 | 43.6 | 578 | 12 | ADI44256 | Adi44256 Plant tra |
| 12 | 286 | 43.6 | 578 | 14 | AEA27034 | Aea27034 Stress to |
| 13 | 285 | 43.4 | 482 | 14 | ADW16587 | Adw16587 Eucalyptu |
| 14 | 277.5 | 42.3 | 504 | 3 | AAC49819 | Aac49819 Arabidops |
| 15 | 276 | 42.1 | 466 | 12 | ADJ43786 | Adj43786 Plant cDN |
| 16 | 275.5 | 42.0 | 294 | 3 | AAC43000 | Aac43000 Arabidops |
| 17 | 275.5 | 42.0 | 294 | 14 | AEA27038 | Aea27038 Stress to |
| 18 | 275.5 | 42.0 | 543 | 13 | ACN61386 | Acn61386 Cotton gy |
| 19 | 275.5 | 42.0 | 779 | 13 | ADR59490 | Adr59490 Cotton cD |
| 20 | 273 | 41.6 | 303 | 3 | AAC46353 | Aac46353 Arabidops |
| 21 | 270.5 | 41.2 | 507 | 13 | ADX14446 | Adx14446 Plant ful |
| 22 | 266 | 40.5 | 636 | 13 | ADR65153 | Adr65153 Cotton cD |
| 23 | 262 | 39.9 | 513 | 11 | ACL30919 | Acl30919 Rice abio |
| 24 | 259.5 | 39.6 | 505 | 13 | ACN61224 | Acn61224 Cotton gy |
| 25 | 259.5 | 39.6 | 532 | 13 | ACN59701 | Acn59701 Cotton gy |
| 26 | 259.5 | 39.6 | 538 | 13 | ACN56695 | Acn56695 Cotton gy |
| 27 | 259.5 | 39.6 | 604 | 13 | ADR64219 | Adr64219 Cotton cD |
| 28 | 259.5 | 39.6 | 604 | 13 | ACN56782 | Acn56782 Cotton gy |
| 29 | 249 | 38.0 | 577 | 13 | ACN59055 | Acn59055 Cotton gy |
| 30 | 245 | 37.3 | 513 | 10 | ADB31830 | Adb31830 DNA encod |
| 31 | 245 | 37.3 | 513 | 12 | ADO02248 | Ado02248 Thalecres |
| 32 | 245 | 37.3 | 513 | 14 | AEA27102 | Aea27102 Stress to |
| 33 | 239 | 36.4 | 776 | 12 | ADJ39433 | Adj39433 Plant cDN |
| 34 | 233.5 | 35.6 | 980 | 13 | ADX51987 | Adx51987 Plant ful |
| 35 | 231 | 35.2 | 2192 | 3 | AAA78533 | Aaa78533 Plant SDF |
| 36 | 228.5 | 34.8 | 584 | 13 | ACN54517 | Acn54517 Cotton an |
| 37 | 218 | 33.2 | 2106 | 14 | ADW16604 | Adw16604 Eucalyptu |
| 38 | 215 | 32.8 | 1943 | 13 | ADW16980 | Adw16980 Pinus rad |
| 39 | 214.5 | 32.7 | 1136 | 13 | ADX12817 | Adx12817 Plant ful |
| 40 | 212.5 | 32.4 | 661 | 12 | ADO62761 | Ado62761 Transcrip |
| 41 | 211 | 32.2 | 866 | 6 | ABK65370 | Abk65370 Arabidops |
| 42 | 211 | 32.2 | 866 | 10 | ADD30950 | Add30950 Plant yle |
| 43 | 211 | 32.2 | 866 | 10 | ADE37148 | Ade37148 Plant yle |
| 44 | 211 | 32.2 | 866 | 12 | ADI41988 | Adi41988 Plant tra |
| 45 | 211 | 32.2 | 866 | 12 | ADO03516 | Ado03516 Thalecres |

ALIGNMENTS

RESULT 1
ADP43565

ID ADF43565 standard; DNA; 381 BP.

AC ADF43565;

DT 12-FEB-2004 (first entry)

DE Thale cress peronospora parasitica resistant protein coding sequence.

KW transgenic plant; peronospora parasitica resistant; PPR2;

KW pathogen resistant phenotype; modified pathogen resistance; thale cress;
KW gene; ds.

OS Arabidopsis thaliana.

PH Key Location/Qualifiers

FT CDS 1..381

FT /*tag= a

FT /product= "Thale cress peronospora parasitica resistant

protein"

XX WO2003091412-A2.

XX PD 06-NOV-2003.

XX PF 24-APR-2003; 2003WO-US012981.

XX

```
PR 24-APR-2002; 2002US-0375333P.
XX
XX (AGRI-) AGRINOMICS LLC.
XX
XX Federpiel N, Lammers A, Liu XL, Bates SR, Westerlund C;
PI Fitch JR;
XX
XX WPI; 2003-865582/80.
DR P-PSDB; ADF43566.
XX
XX New transgenic plants with increased resistance to pathogens due to
PT altered expression of Peronospora Parasitica Resistant gene (PPR2),
PT useful for generating plants with a pathogen resistance phenotype.
XX
XX Example 4; SEQ ID NO 1; 36pp; English.
XX
XX The invention relates to a transgenic plant that possesses a plant
CC transformation vector comprising a nucleotide sequence that encodes a
CC peronospora parasitica resistant (PPR2) protein, or a PPR2 orthologue.
CC The transgenic plant is useful in generating plants with a pathogen
CC resistance phenotype. The PPR2 nucleic acids and proteins of the
CC invention are useful in the generation of genetically modified plants
CC having a modified pathogen resistance phenotype. The present DNA sequence
CC encodes an Arabidopsis thaliana PPR2 protein of the invention.
XX
XX Sequence 381 BP; 122 A; 76 C; 76 G; 107 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4.3e-76 Length: 381
Score: 656.00 Matches: 126
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-697-787-2 (1-126) x ADF43565 (1-381)
Qy 1 MetAlaSerAsnSerArgSerSeriLeSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAGAGTTCAATCTCACCATTGGACGTTTAGTCAAAACAAGATGTTTC 60
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAGGACACACCCGCGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspLysLeuValGluAsp 60
Db 121 GCTGTCGGAGGAAACTGTAGAGAGTGTAGAGCCCACTATGACATTCCTCGTCGAGGAT 180
Qy 61 LeuileAsnLeuGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGGCCCAATTACAAGACCTTCGAATCTAAC 240
Qy 81 SerArgSerileAsnAspPheAspThrArgTyrileThrLysTyrLeuTyrMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGACACAGGTATATACTAAATATCTATATATGATGCTC 300
Qy 101 SerileTyrPheAspAsnHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATATATCATCTTAGTGATTTTGAGAAATCTCTCAAAAGTCTTGTA 360
Qy 121 SerTyrileSerLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378
RESULT 2
AEA271134
ID AEA271134 standard; DNA; 294 BP.
XX
XX AEA271134;
XX
XX 28-JUL-2005 (first entry)
XX
```

```
DE Stress tolerant plant-related transcription factor gene SeqID975.
XX
XX transcription factor; transgenic plant; agriculture; drought resistance;
KW stress tolerance; gene; ds.
XX
XX Arabidopsis thaliana.
OS
XX W02005047516-A2.
PN
XX 26-MAY-2005.
PD
XX
XX 12-NOV-2004; 2004WO-US037584.
PF
XX
XX 13-NOV-2003; 2003US-00714887.
PR
XX 05-DEC-2003; 2003US-0527658P.
PR
XX 05-FEB-2004; 2004US-0542928P.
PR
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA
XX
XX Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kumimoto RW, Gutterman NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Keddle JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JM;
XX
XX WPI; 2005-372386/38.
DR P-PSDB; AEA271135.
XX
XX New transgenic plants for producing commercially or agriculturally useful
PT plants having improved tolerance to drought, shade and low nitrogen
PT conditions.
XX
XX Disclosure; SEQ ID NO 975; 407pp; English.
XX
XX This invention relates to a novel plant transcription factor
CC polypeptides, the DNA sequences which encode them and their use in
CC creating transgenic plants. The transgenic plant and methods are useful
CC for producing commercially or agriculturally useful plants having
CC improved tolerance to drought, shade and low nitrogen conditions when
CC compared to wild-type reference plants. The present sequence is that of a
CC plant transcription factor gene which was used during the development of
CC the transgenic plants of the invention.
XX
XX Sequence 294 BP; 99 A; 65 C; 67 G; 63 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2.39e-52 Length: 294
Score: 473.00 Matches: 90
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 72.1% Indels: 0
DB: 14 Gaps: 0
US-10-697-787-2 (1-126) x AEA271134 (1-294)
Qy 1 MetAlaSerAsnSerArgSerSeriLeSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATTGGACGTTTAGTCAAAACAAGATGTTTC 60
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAGGACACACCCGCGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspLysLeuValGluAsp 60
Db 121 GCTGTCGGAGGAAACTGTAGAGAGTGTAGAGCCCACTATGACATTCCTCGTCGAGGAT 180
Qy 61 LeuileAsnLeuGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGGCCCAATTACAAGACCTTCGAATCTAAC 240
Qy 81 SerArgSerileAsnAspPheAspThrArg 90
Db 241 TCAAGAAGCATCAATGACTTTTGACACAGG 270
```

RESULT 3
AD062042
ID AD062042 standard; DNA; 279 BP.
XX
AC AD062042;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2723 coding sequence, SEQ ID 509.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
DR WPI: 2004-330163/30.
DR P-PSDB; AD062043.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 509; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil

CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 279 BP; 86 A; 56 C; 65 G; 72 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.11e-35 Length: 279
Score: 343.00 Matches: 69
Percent Similarity: 80.4% Conservative: 5
Best Local Similarity: 75.0% Mismatches: 16
Query Match: 52.3% Indels: 2
DB: 12 Gaps: 1
US-10-697-787-2 (1-126) x ADO62042 (1-279)
QY 1 MetAlaSerAsnSerArgSerIleSerProTTPThrPheSerGlnAsnLysMetPhe 20
DB 1 ATGGCTCTAGTTCATGAGCTCTTCTTGACGCTCTAAGCAAAACGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHisAsnValAlaLys 40
DB 61 GAGAGGGCTTTAGCCGTTTACGATAAGACACTCCGACCGTTGGCAAAACGTTGCTAAA 120
QY 41 AlaValGlyGlyThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
DB 121 GCAGTTGGAAAGTAAATCTGCAGAGGAAGTTAAACGCTCACTAGCACATCTCGTTGAAGAT 180
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu----- 78
DB 181 CTCATGAACATCGAACAAAGACTTAGTACCTTTGGCTTAATAACAAACCGTCGATGTGGA 240
QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArg 90
DB 241 AGTAAATCTAGAGGCGATCGATGATTTCGATTGAGG 276
RESULT 4
AAC34192
ID AAC34192 standard; DNA; 478 BP.
XX
AC AAC34192;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5783.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.

| | | |
|----|--------------|-----------------|
| PR | 23-APR-1999; | 99US-0130510P. |
| PR | 23-APR-1999; | 99US-0130891P. |
| PR | 28-APR-1999; | 99US-01311449P. |
| PR | 30-APR-1999; | 99US-0132048P. |
| PR | 30-APR-1999; | 99US-0132407P. |
| PR | 04-MAY-1999; | 99US-0132484P. |
| PR | 05-MAY-1999; | 99US-0132485P. |
| PR | 06-MAY-1999; | 99US-0132486P. |
| PR | 06-MAY-1999; | 99US-0132487P. |
| PR | 07-MAY-1999; | 99US-0132863P. |
| PR | 11-MAY-1999; | 99US-0134256P. |
| PR | 14-MAY-1999; | 99US-0134218P. |
| PR | 14-MAY-1999; | 99US-0134219P. |
| PR | 14-MAY-1999; | 99US-0134221P. |
| PR | 14-MAY-1999; | 99US-0134370P. |
| PR | 16-MAY-1999; | 99US-0134768P. |
| PR | 19-MAY-1999; | 99US-0134941P. |
| PR | 20-MAY-1999; | 99US-0135124P. |
| PR | 21-MAY-1999; | 99US-0135353P. |
| PR | 24-MAY-1999; | 99US-0135629P. |
| PR | 25-MAY-1999; | 99US-0136021P. |
| PR | 27-MAY-1999; | 99US-0136332P. |
| PR | 28-MAY-1999; | 99US-0136782P. |
| PR | 01-JUN-1999; | 99US-0137222P. |
| PR | 03-JUN-1999; | 99US-0137528P. |
| PR | 04-JUN-1999; | 99US-0137502P. |
| PR | 07-JUN-1999; | 99US-0137724P. |
| PR | 08-JUN-1999; | 99US-0138094P. |
| PR | 10-JUN-1999; | 99US-0138540P. |
| PR | 10-JUN-1999; | 99US-0138847P. |
| PR | 14-JUN-1999; | 99US-0139119P. |
| PR | 16-JUN-1999; | 99US-0139452P. |
| PR | 16-JUN-1999; | 99US-0139453P. |
| PR | 17-JUN-1999; | 99US-0139492P. |
| PR | 18-JUN-1999; | 99US-0139454P. |
| PR | 18-JUN-1999; | 99US-0139455P. |
| PR | 18-JUN-1999; | 99US-0139456P. |
| PR | 18-JUN-1999; | 99US-0139457P. |
| PR | 18-JUN-1999; | 99US-0139458P. |
| PR | 18-JUN-1999; | 99US-0139459P. |
| PR | 18-JUN-1999; | 99US-0139460P. |
| PR | 18-JUN-1999; | 99US-0139461P. |
| PR | 18-JUN-1999; | 99US-0139462P. |
| PR | 18-JUN-1999; | 99US-0139463P. |
| PR | 18-JUN-1999; | 99US-0139750P. |
| PR | 18-JUN-1999; | 99US-0139763P. |
| PR | 21-JUN-1999; | 99US-0139817P. |
| PR | 22-JUN-1999; | 99US-0139899P. |
| PR | 23-JUN-1999; | 99US-0140353P. |
| PR | 23-JUN-1999; | 99US-0140354P. |
| PR | 24-JUN-1999; | 99US-0140695P. |
| PR | 28-JUN-1999; | 99US-0140823P. |
| PR | 29-JUN-1999; | 99US-0140991P. |
| PR | 30-JUN-1999; | 99US-0141287P. |
| PR | 01-JUL-1999; | 99US-0141842P. |
| PR | 01-JUL-1999; | 99US-0142154P. |
| PR | 02-JUL-1999; | 99US-0142055P. |
| PR | 06-JUL-1999; | 99US-0142300P. |
| PR | 08-JUL-1999; | 99US-0142803P. |
| PR | 09-JUL-1999; | 99US-0142920P. |
| PR | 12-JUL-1999; | 99US-0142977P. |
| PR | 13-JUL-1999; | 99US-0143542P. |
| PR | 14-JUL-1999; | 99US-0143624P. |
| PR | 15-JUL-1999; | 99US-0144005P. |
| PR | 16-JUL-1999; | 99US-0144085P. |
| PR | 16-JUL-1999; | 99US-0144086P. |
| PR | 19-JUL-1999; | 99US-0144325P. |
| PR | 19-JUL-1999; | 99US-0144331P. |
| PR | 19-JUL-1999; | 99US-0144332P. |
| PR | 19-JUL-1999; | 99US-0144333P. |
| PR | 19-JUL-1999; | 99US-0144334P. |
| PR | 19-JUL-1999; | 99US-0144335P. |
| PR | 20-JUL-1999; | 99US-0144352P. |
| PR | 20-JUL-1999; | 99US-0130510P. |
| PR | 20-JUL-1999; | 99US-0130891P. |
| PR | 21-JUL-1999; | 99US-01311449P. |
| PR | 21-JUL-1999; | 99US-0132048P. |
| PR | 21-JUL-1999; | 99US-0132407P. |
| PR | 22-JUL-1999; | 99US-0132484P. |
| PR | 22-JUL-1999; | 99US-0132485P. |
| PR | 22-JUL-1999; | 99US-0132486P. |
| PR | 22-JUL-1999; | 99US-0132487P. |
| PR | 23-JUL-1999; | 99US-0132863P. |
| PR | 23-JUL-1999; | 99US-0134256P. |
| PR | 23-JUL-1999; | 99US-0134218P. |
| PR | 26-JUL-1999; | 99US-0134219P. |
| PR | 27-JUL-1999; | 99US-0134221P. |
| PR | 27-JUL-1999; | 99US-0134370P. |
| PR | 28-JUL-1999; | 99US-0134768P. |
| PR | 02-AUG-1999; | 99US-0134941P. |
| PR | 02-AUG-1999; | 99US-0135124P. |
| PR | 02-AUG-1999; | 99US-0135353P. |
| PR | 03-AUG-1999; | 99US-0135629P. |
| PR | 04-AUG-1999; | 99US-0136021P. |
| PR | 04-AUG-1999; | 99US-0136332P. |
| PR | 05-AUG-1999; | 99US-0136782P. |
| PR | 05-AUG-1999; | 99US-0137222P. |
| PR | 06-AUG-1999; | 99US-0137528P. |
| PR | 06-AUG-1999; | 99US-0137502P. |
| PR | 09-AUG-1999; | 99US-0137724P. |
| PR | 09-AUG-1999; | 99US-0138094P. |
| PR | 10-AUG-1999; | 99US-0138540P. |
| PR | 11-AUG-1999; | 99US-0138847P. |
| PR | 12-AUG-1999; | 99US-0139119P. |
| PR | 13-AUG-1999; | 99US-0139452P. |
| PR | 13-AUG-1999; | 99US-0139453P. |
| PR | 16-AUG-1999; | 99US-0139492P. |
| PR | 17-AUG-1999; | 99US-0139454P. |
| PR | 18-AUG-1999; | 99US-0139455P. |
| PR | 18-AUG-1999; | 99US-0139456P. |
| PR | 18-AUG-1999; | 99US-0139457P. |
| PR | 20-AUG-1999; | 99US-0149722P. |
| PR | 20-AUG-1999; | 99US-0149723P. |
| PR | 23-AUG-1999; | 99US-0149902P. |
| PR | 23-AUG-1999; | 99US-0149910P. |
| PR | 25-AUG-1999; | 99US-0150566P. |
| PR | 26-AUG-1999; | 99US-0150884P. |
| PR | 27-AUG-1999; | 99US-0151065P. |
| PR | 27-AUG-1999; | 99US-0151066P. |
| PR | 30-AUG-1999; | 99US-0151080P. |
| PR | 31-AUG-1999; | 99US-0151303P. |
| PR | 01-SEP-1999; | 99US-0151438P. |
| PR | 07-SEP-1999; | 99US-0151930P. |
| PR | 10-SEP-1999; | 99US-0152363P. |
| PR | 13-SEP-1999; | 99US-0153070P. |
| PR | 15-SEP-1999; | 99US-0153758P. |
| PR | 16-SEP-1999; | 99US-0154018P. |
| PR | 20-SEP-1999; | 99US-0154039P. |
| PR | 22-SEP-1999; | 99US-0154779P. |
| PR | 23-SEP-1999; | 99US-0155139P. |
| PR | 24-SEP-1999; | 99US-0155486P. |
| PR | 24-SEP-1999; | 99US-0155659P. |
| PR | 28-SEP-1999; | 99US-0156458P. |
| PR | 29-SEP-1999; | 99US-0156596P. |
| PR | 04-OCT-1999; | 99US-0157117P. |
| PR | 05-OCT-1999; | 99US-0157753P. |
| PR | 06-OCT-1999; | 99US-0157865P. |
| PR | 07-OCT-1999; | 99US-0158029P. |
| PR | 08-OCT-1999; | 99US-0158232P. |
| PR | 12-OCT-1999; | 99US-0158369P. |
| PR | 13-OCT-1999; | 99US-0159293P. |
| PR | 13-OCT-1999; | 99US-0159294P. |
| PR | 13-OCT-1999; | 99US-0159295P. |
| PR | 14-OCT-1999; | 99US-0159329P. |
| PR | 14-OCT-1999; | 99US-0159330P. |
| PR | 14-OCT-1999; | 99US-0159331P. |

PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161320P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
 Pred. No.: 4,64e-33 Length: 478
 Score: 327.50 Matches: 67
 Percent Similarity: 73.2% Conservatives: 4
 Best Local Similarity: 69.1% Mismatches: 19
 Query Match: 49.9% Indels: 7
 DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x AAC34192 (1-478)

QY 10 SerProThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLys 29
 Db 5 TCTTCTTGACGCTTAAGCAAAACAGATGTTTCGAGAGGCGTTTACGCTTTACGATAAA 64
 QY 30 AspThrProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGlu 49
 Db 65 GACACTCCGCGCGTGGCAAAACGTCGCTTAAGCAGTTGGAAGTAATCTCCAGAGNA 124
 QY 50 ValLysArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgVal 69
 Db 125 GTTAAACGTCCTACGACATCTCGTTGAAGATCTCATGAACATCGAACAGACTTAGTA 184
 QY 70 ProLeuProAsnTyrLysThrPheGlu-----SerAsnSerArgSerIleAsnAspPhe 87
 Db 185 CTTTTCCTTAAATACAAACCGTCGATGTTGGAAGTAATCTAGAGGAATCAATGATTTT 244
 QY 88 AspThrArgTyrIleThrLysTyr-----LeuTyrMetMet 99
 Db 245 GATTTGAGGTTATTGAGAAATATGAGAAATCCAGTCCAGTGCATCATGATG 295

RESULT 5

ID ADR59489 standard; cDNA; 760 BP.

XX AC ADR59489;

XX DT 02-DEC-2004 (first entry)

XX DE Cotton cDNA sequence, SEQ ID 270.

KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KW drought tolerance; plant disease resistance; galactomannan; lignin;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW homologous recombination; extreme osmotic condition tolerance;
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 KW stress resistance.

OS Gossypium hirsutum.

XX US2004181030-A1.

QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60

XX 16-SBP-2004.
 XX 29-JAN-2004; 2004US-00767795.
 XX 07-MAY-2001; 2001US-00849529.
 PR 12-DEC-2001; 2001US-00021323.
 XX (KOVA/) KOVALIC D K.
 PA (ZHOU/) ZHOU Y.
 PA (CAOY/) CAO Y.
 XX Kovalic DK, Zhou Y, Cao Y;
 XX WPI; 2004-667718/65.
 XX New recombinant nucleic acid molecules and polypeptides from Gossypium
 PT hirsutum, useful for producing plants with improved biological
 PT characteristics (e.g. improved plant cold or drought tolerance).
 XX Claim 1; SEQ ID NO 270; 14pp; English.

XX The invention relates to a recombinant polynucleotide comprising any of
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid
 CC sequences mentioned in the specification and producing a plant having an
 CC improved property. Producing a plant having an improved property
 CC comprises transforming a plant with a recombinant construct comprising a
 CC promoter region functional in a plant cell operably joined to a
 CC polynucleotide comprising a coding sequence for a polypeptide associated
 CC with the property, and growing the transformed plant. The polypeptide is
 CC useful for improving plant cold tolerance, manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, improving plant
 CC drought tolerance, providing increased resistance to plant disease,
 CC producing galactomannan (or lignin or plant growth regulators), improving
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing
 CC the rate of homologous recombination in plants, improving plant tolerance
 CC to extreme osmotic conditions or to pathogens or pests, improving yield
 CC by modification of photosynthesis, modifying seed oil or protein yield
 CC or phosphorus use and/or uptake, or improving yield by providing improved
 CC plant growth and development under at least one stress condition. The
 CC polynucleotide and polypeptide may also be used in recombinant DNA
 CC constructs, in physical arrays of molecules, as plant breeding markers,
 CC or in computer-based storage and analysis systems. The present sequence,
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040181030. However only 6585
 CC polynucleotide sequences were available, the remaining 52213
 CC polynucleotides and all 58798 protein sequences were not present.

XX Sequence 760 BP; 308 A; 90 C; 144 G; 218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,9e-28 Length: 760
 Score: 290.00 Matches: 56
 Percent Similarity: 83.8% Conservatives: 11
 Best Local Similarity: 70.0% Mismatches: 11
 Query Match: 44.2% Indels: 2
 DB: 13 Gaps: 1

US-10-697-787-2 (1-126) x ADR59489 (1-760)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProThrPheSerGlnAsnLysMetPhe 20
 Db 113 ATGTCATCGAATTCATGCTCT-----GCTTCATGGACAGCAACAAACAGATTTC 166
 QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValLys 40
 Db 167 GAAAGGGCTTTAGCTGTTTACGACAAAGGACACACAGATCGTTGGTACAAATGTTCTAAA 226

```
Db      227 GCTGTGGGAGGAAACATGTTGAGGAAGTGAAGACCATATGAGCTTCTCTGGAAGAT 286
QY      61 LeuileAsnileGluThrGlyArgValProLeuProAsnTyrLyThrPheGluSerAsn 80
Db      287 GTTAGACACATCGATCGGTCGGTTCCTTCCCCGACTATTGGACCGTTACCGGGAAT 346

RESULT 6
AAC35256
ID AAC35256 standard; DNA; 508 BP.
AC AAC35256;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9536.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 08-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130851P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147453P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
```


CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX
SQ Sequence 704 BP; 202 A; 159 C; 121 G; 222 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1.53e-27 | Length: | 704 |
| Score: | 287.00 | Matches: | 54 |
| Percent Similarity: | 80.2% | Conservative: | 11 |
| Best Local Similarity: | 66.7% | Mismatches: | 16 |
| Query Match: | 43.8% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |

US-10-697-787-2 (1-126) x ADX32012 (1-704)

| | | | |
|----|-----|---|-----|
| Qy | 1 | MetAlaSerAsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPhe | 20 |
| Db | 150 | ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTGTAAAGGACAAAGGCCCTTT | 209 |
| Qy | 21 | GluArgAlaLeuAlaValTyrAspLysPheThrProAspArgTrrPheHisAsnValAlaLys | 40 |
| Db | 210 | GAAGAAGCTTAGCTGTTTATGACAAAGGACACTCTGACCGTTGGTACATGTGTCAT | 269 |
| Qy | 41 | AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp | 60 |
| Db | 270 | GCTGTTGGTGGCAAAACTCCAGAGGAAGTCAAAAGGCACTACGAACTCCTTGTTCAGCAT | 329 |
| Qy | 61 | LeuIleAsnIleGluThrGlyArgValProLeuProAsnTrrLysThrPheGluSerAsn | 80 |
| Db | 330 | GTTAAGCATATTGAGTCTGGACGTGTGCCATTCCCAAAATTACAGAAAACTACTTTCAGGG | 389 |
| Qy | 81 | Ser | 81 |
| Db | 390 | TCA | 392 |

RESULT 8

AAC42715

ID AAC42715 standard; DNA; 306 BP.

XX

AC AAC42715;

XX

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 36586.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX

PD 06-SEP-2000.

XX

XX 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

| | | |
|----|--------------|----------------|
| PR | 08-APR-1999; | 99US-0128714P. |
| PR | 16-APR-1999; | 99US-0129645P. |
| PR | 19-APR-1999; | 99US-0130077P. |
| PR | 21-APR-1999; | 99US-0130449P. |
| PR | 23-APR-1999; | 99US-0130510P. |
| PR | 23-APR-1999; | 99US-0130891P. |
| PR | 28-APR-1999; | 99US-0131449P. |
| PR | 30-APR-1999; | 99US-0132048P. |
| PR | 30-APR-1999; | 99US-0132407P. |
| PR | 04-MAY-1999; | 99US-0132484P. |
| PR | 05-MAY-1999; | 99US-0132486P. |
| PR | 06-MAY-1999; | 99US-0132487P. |
| PR | 07-MAY-1999; | 99US-0132863P. |
| PR | 11-MAY-1999; | 99US-0134256P. |
| PR | 14-MAY-1999; | 99US-0134218P. |
| PR | 14-MAY-1999; | 99US-0134219P. |
| PR | 14-MAY-1999; | 99US-0134221P. |
| PR | 14-MAY-1999; | 99US-0134370P. |
| PR | 18-MAY-1999; | 99US-0134768P. |
| PR | 19-MAY-1999; | 99US-0134941P. |
| PR | 20-MAY-1999; | 99US-0135124P. |
| PR | 21-MAY-1999; | 99US-0135353P. |
| PR | 24-MAY-1999; | 99US-0135629P. |
| PR | 25-MAY-1999; | 99US-0136021P. |
| PR | 27-MAY-1999; | 99US-0136392P. |
| PR | 28-MAY-1999; | 99US-0136782P. |
| PR | 01-JUN-1999; | 99US-0137222P. |
| PR | 03-JUN-1999; | 99US-0137528P. |
| PR | 04-JUN-1999; | 99US-0137502P. |
| PR | 07-JUN-1999; | 99US-0137724P. |
| PR | 08-JUN-1999; | 99US-0138094P. |
| PR | 10-JUN-1999; | 99US-0138540P. |
| PR | 10-JUN-1999; | 99US-0138847P. |
| PR | 14-JUN-1999; | 99US-0139119P. |
| PR | 16-JUN-1999; | 99US-0139452P. |
| PR | 16-JUN-1999; | 99US-0139453P. |
| PR | 17-JUN-1999; | 99US-0139492P. |
| PR | 18-JUN-1999; | 99US-0139454P. |
| PR | 18-JUN-1999; | 99US-0139455P. |
| PR | 18-JUN-1999; | 99US-0139456P. |
| PR | 18-JUN-1999; | 99US-0139457P. |
| PR | 18-JUN-1999; | 99US-0139458P. |
| PR | 18-JUN-1999; | 99US-0139459P. |
| PR | 18-JUN-1999; | 99US-0139460P. |
| PR | 18-JUN-1999; | 99US-0139461P. |
| PR | 18-JUN-1999; | 99US-0139462P. |
| PR | 18-JUN-1999; | 99US-0139463P. |
| PR | 18-JUN-1999; | 99US-0139750P. |
| PR | 18-JUN-1999; | 99US-0139763P. |
| PR | 21-JUN-1999; | 99US-0139817P. |
| PR | 22-JUN-1999; | 99US-0139899P. |
| PR | 23-JUN-1999; | 99US-0140353P. |
| PR | 23-JUN-1999; | 99US-0140354P. |
| PR | 24-JUN-1999; | 99US-0140695P. |
| PR | 28-JUN-1999; | 99US-0140823P. |
| PR | 29-JUN-1999; | 99US-0140991P. |
| PR | 30-JUN-1999; | 99US-0141287P. |
| PR | 01-JUL-1999; | 99US-0141842P. |
| PR | 01-JUL-1999; | 99US-0142154P. |
| PR | 02-JUL-1999; | 99US-0142055P. |
| PR | 06-JUL-1999; | 99US-0142390P. |
| PR | 08-JUL-1999; | 99US-0142803P. |
| PR | 09-JUL-1999; | 99US-0142920P. |
| PR | 12-JUL-1999; | 99US-0142977P. |
| PR | 13-JUL-1999; | 99US-0143542P. |
| PR | 14-JUL-1999; | 99US-0143624P. |
| PR | 15-JUL-1999; | 99US-0144005P. |
| PR | 16-JUL-1999; | 99US-0144085P. |
| PR | 16-JUL-1999; | 99US-0144086P. |
| PR | 19-JUL-1999; | 99US-0144325P. |
| PR | 19-JUL-1999; | 99US-0144331P. |
| PR | 19-JUL-1999; | 99US-0144332P. |

PN US2002040489-A1.
XX 04-APR-2002.
PD 26-JAN-2001; 2001US-00770152.
XX 27-JAN-2000; 2000US-0178503P.
XX (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2003-110410/10.
XX
PT Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 473; 45pp; English.
XX
CC The invention relates to Arabidopsis thaliana nucleic acid sequences. The
CC DNA sequences and the polypeptides they encode are useful for identifying
CC homologous or related genes, for producing compositions that modulate the
CC expression or function of the polypeptides, for mapping functional
CC regions of the protein, in diagnosis, for studying associated
CC physiological pathways, for genetic manipulation of cells, preferably
CC plant cells, in screening assays of various plant strains to determine
CC the strains that are capable of withstanding a particular disease or
CC environmental stress, for enhancing or inhibiting production of
CC biosynthetic products in plants and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. Transgenic plants
CC are useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biologically active agents, such as
CC fungicides and insecticides, and for elucidating biochemical pathways.
CC Sequences ABX5649-ABX57647 represent Arabidopsis thaliana
CC polynucleotides of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 556 BP; 146 A; 114 C; 115 G; 181 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.5e-27 Length: 556
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 10 Gaps: 1
US-10-697-787-2 (1-126) x ABX57121 (1-556)
Qy 1 MetAlaSerAsnSerArgSerSerIle-----SerProTrpThrPheSerGlnAsnLys 18

Db 474 ATGGCATCAGGCTCAATGTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 415
Qy 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
Db 414 GCCTTTGAGGCTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 355
Qy 39 AlaLysAlaValGlyGlyThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Db 354 GCTAGAGCTGTTGGTGGTAAACACCCAGAAGAAGCTTAAGAGACAGTATGACCTTCTAGTT 295
Qy 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
Db 294 CGTGACATCGAAGCATCGAAGATGGTCACGGTCCATTCCTGACTACAAGACTACTACA 235
Qy 79 SerAsnSer 81
Db 234 GGAACACAGC 226
RESULT 10
ADD30327
ID ADD30327 standard; cDNA; 578 BP.
XX
AC ADD30327;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1789.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
PN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
(MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Broun PE;
XX
DR WPI; 2003-248221/24.
DR P-PSDB; ADD30328.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 356; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcribed
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.

```
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.58e-27 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 12 Gaps: 1

US-10-697-787-2 (1-126) x ADD30327 (1-578)
QY 1 MetAlaSerAsnSerArgSerSerIle-----SerProTrrPrrPheSerGlnAsnLys 18
DB 108 ATGGCATCAGGCTCAATGCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 167
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 168 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 227
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
DB 228 GCTAGAGCTGTGGTGTGTAACACACAGAGAGCTAAGAGACAGTATGACCTTCTAGTT 287
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
DB 288 CGTGACATCGAAAGCATCGAGAAATGTCACGTGCCATTCCTGACTACAGACTACTACA 347
QY 79 SerAsnSer 81
DB 348 GGAACACAGC 356

RESULT 11
ADI44256
ID ADI44256 standard; DNA; 578 BP.
XX AC ADI44256;
XX 22-APR-2004 (first entry)
XX DE Plant transcription factor related polynucleotide #1717.
XX KW transgenic; plant; enhanced tolerance to abiotic stress;
KW glyophosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; ds.
XX OS Unidentified.
XX PN US2004019927-A1.
XX PD 29-JAN-2004.
XX PF 25-FEB-2003; 2003US-00374780.
XX PR 18-APR-2001; 2001US-00837944.
XX PA (SHER/) SHERMAN B K.
PA (RIE/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAR/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.

(BROU/) BROUN P E.
(PILG/) PILGRIM M L.
(DUBE/) DUBELL A N.
(PINE/) PINEDA O.
(YUGG/) YU G.
Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V,
Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PB;
Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
DR P-PSDB; ADI44257.
XX New transgenic plant comprising a recombinant polynucleotide of any one
of more than 500 nucleotide sequences, useful in bioinformatic search
methods.
XX Disclosure; SEQ ID NO 2719; 435pp; English.
XX The invention describes a transgenic plant comprising a recombinant
polynucleotide of any one of more than 500 nucleotide sequences fully
defined in the specification or its complement. The method of the
invention can be used to produce a plant having altered traits such as:
enhanced tolerance to abiotic stress; glyophosphate tolerance; hormone
sensitivity; disease resistance; sugar sensing; early or late flowering;
altered flower structure, change in stem bifurcations, altered branching
pattern, reduced apical dominance, reduced trichome density; lack of
trichomes; reduced ectopic trichome development; altered trichome
development; increase in trichome number; altered stem morphology;
increased root growth; increased root hairs; altered seed development;
altered cell proliferation or cell differentiation; rapid development;
premature senescence; increased necrosis; increase in seedling or plant
size; decreased plant size; leaf morphology; seed morphology; seed
biochemistry; increase in root anthocyanins; increase in plant
anthocyanins, or alteration in light response or shade avoidance. The
transgenic plant, polynucleotides and polypeptides are useful in
bioinformatic search methods. This sequence represents a plant
transcription factor related polynucleotide.
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;
```

ID AEA27034 standard; DNA; 578 BP.
XX AC AEA27034;
XX DT 28-JUL-2005 (first entry)
XX DE Stress tolerant plant-related transcription factor gene SeqID875.
XX KW transcription factor; transgenic plant; agriculture; drought resistance;
XX KW stress tolerance; gene; ds.
XX OS Arabidopsis thaliana.
XX PN WO2005047516-A2.
XX PD 26-MAY-2005.
XX PF 12-NOV-2004; 2004WO-US037584.
XX PR 13-NOV-2003; 2003US-00714887.
XX PR 05-DEC-2003; 2003US-0527658P.
XX PR 05-FEB-2004; 2004US-0542928P.
XX PA (WEND-) MENDEL BIOTECHNOLOGY INC.
XX PI Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kumimoto RW, Guttererson NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Keddle JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JM;
XX PS WPI; 2005-372386/38.
XX DR P-PSDB; AEA27035.
XX PT New transgenic plants for producing commercially or agriculturally useful
XX PT plants having improved tolerance to drought, shade and low nitrogen
XX PT conditions.
XX PS Disclosure; SEQ ID NO 875; 407pp; English.
XX CC This invention relates to a novel plant transcription factor
XX CC polypeptides, the DNA sequences which encode them and their use in
XX CC creating transgenic plants. The transgenic plant and methods are useful
XX CC for producing commercially or agriculturally useful plants having
XX CC improved tolerance to drought, shade and low nitrogen conditions when
XX CC compared to wild-type reference plants. The present sequence is that of a
XX CC plant transcription factor gene which was used during the development of
XX CC the transgenic plants of the invention.
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.59e-27 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 14 Gaps: 1

US-10-697-787-2 (1-126) x AEA27034 (1-578)

Qy 1 MetAlaSerAenSerArgSerSerIle-----SerProTprThrPheSerGlnAenLys 18
Db 108 ATGCATCAGGCTCAATGTCTTATTATGGCTTCATGGCTCATGGACTGTTAAGCAGACAAA 167
Qy 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAenVal 38
Db 168 GCCTTTGAGGTGCTTAGCAGTCTATGACACACACTCCGGACCTTGGCACAATGTT 227
Qy 39 AlalysAlaValGlyGlyLysThrValGluValLysArgHisTyrAspIleLeuVal 58
Db 228 GCTAGACGTGTGTGGTAAACACCAAGAGACAGTATGACCTTCTAGTT 287
Qy 59 GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAAGCATCGAGAAATGTCACGTGCCATTCCTGACTACAGACTACTACA 347
Qy 79 SerAenSer 81
Db 348 GGAAACAGC 356
RESULT 13
ADW16587
ID ADW16587 standard; cDNA; 482 BP.
XX AC ADW16587;
XX DT 24-MAR-2005 (first entry)
XX DE Eucalyptus grandis transcription factor cDNA MYB family Seq 306.
XX KW gene; ss; plant; transcription; gene regulation; gene expression;
XX KW transgenic plant; drought resistance; disease resistance; salt tolerance;
XX KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
XX KW flower color.
XX OS Eucalyptus grandis.
XX PN WO2005001050-A2.
XX PD 06-JAN-2005.
XX PF 07-JUN-2004; 2004WO-US017965.
XX PR 06-JUN-2003; 2003US-0476189P.
XX PA (ARBO-) ARBORGEN LLC.
XX PI Blokeberg LM, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX PS WPI; 2005-075542/08.
XX DR P-PSDB; ADW17373.
XX PT New polynucleotides isolated from plants encoding transcription factors,
XX PT and polypeptides encoded by such polynucleotides, useful for regulating
XX PT gene transcription and gene expression.
XX PS Claim 3; SEQ ID NO 306; 1265pp; English.
XX CC This invention relates to novel isolated plant nucleic acid molecules, or
XX CC variants thereof, that encode transcription factors. Specifically, it
XX CC refers to transcription factor proteins that are capable of binding to
XX CC DNA in order to regulate gene transcription and gene expression in a
XX CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
XX CC invention describes DNA constructs containing DNA encoding a
XX CC transcription factor that regulates the promoter, which is operably
XX CC linked to the desired nucleic acid to be expressed. It further provides
XX CC transgenic plants expressing a transcription factor that confers a trait
XX CC to the plant such as increased drought, salt or disease tolerance, height
XX CC change, enhanced cold/ frost tolerance, enhanced taste, health and
XX CC nutritional characteristics, as well as improved color, starch
XX CC composition, flower longevity and germination, amongst others.
XX CC Accordingly, such plants that are successfully transfected with a DNA
XX CC construct can be characterized by a difference in flower color, petal or
XX CC leaf shape and size, aroma or plant height. This polynucleotide is a
XX CC plant transcription factor cDNA sequence of the invention.
XX SQ Sequence 482 BP; 137 A; 89 C; 121 G; 135 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.67e-27 Length: 482
Score: 285.00 Matches: 55
Percent Similarity: 84.0% Conservative: 13
Best Local Similarity: 67.9% Mismatches: 11
Query Match: 43.4% Indels: 2

DB: 14 Gaps:

US-10-697-787-2 (1-126) x ADM16587 (1-482)

Oy 1 MetAlaSerAenSerArGserSeriIleSerProTrpThrPheSerGlnAenLysMetPhe 20
||||| :||| :||| :||| :||| :||| :||| :|||
Db 40 ATGCATCGAATTTCTTGACTTCT---TCATCCTGGACGCCGAAGACAAGAATGTTCC 96

Oy 21 GluArgAlaLeuAlaValTyAspLyAspThrProAspArgTrpHisAenValAlaLys 40
::: ||| :||| :||| :||| :||| :||| :||| :|||
Db 97 GAGAAGGCAATTGGCTCAATATGACAAGCACATCCCACAGGTGGCAGAAGATGCCAAG 156

Oy 41 AlaValGlyGlyLysThrValGlulValLysArgHisiTyraepIleLeuValGluAsp 60
||||| :||| :||| :||| :||| :||| :||| :|||
Db 157 GCCCTGGTGGGAAATCTGCAGATGAAGTGAATAAGACACTATGAATTTTTTAATCGAGGAC 216

Oy 61 LeutlleAenIleGluThrGliVatqValProLeuProAentYrLysThrPheGluSerAan 80
::: ::||| :||| :||| :||| :||| :||| :||| :|||
Db 217 GTCAAGGACATCGAGTCTGGCAGAGTTCTTTTCTTAATTACAGGTGC---AGCAA CAAT 273

Oy 81 Ser 81
|||
Db 274 AGC 276

RESULT 14
ID AAC49819 standard; DNA; 504 BP.
XX AAC49819;
XX AC
XX AC
DT 18-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62553.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-01221825P.
PR 05-MAR-1999; 99US-01233180P.
PR 09-MAR-1999; 99US-01233548P.
PR 23-MAR-1999; 99US-01257888P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130049P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.

```
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151130P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151530P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154019P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.69e-26 Length: 504
Score: 277.50 Matches: 59
Percent Similarity: 69.8% Conservative: 22
Best Local Similarity: 50.9% Mismatches: 25
Query Match: 42.3% Indels: 10
DB: Gaps: 2

US-10-697-787-2 (1-126) x AAC49819 (1-504)

QY 1 MetAlaSerAsnSerArgSerIleSerProTTrpThrPheSerGlnAenLysMetPhe 20
    |||||
Db 103 ATGGCTTCCAACCTCAATGAGCTCTAGCGCTTCTTGACACAGTAAGGAGAACAAATATTT 162
    |||||
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHisAenValAlaLys 40
    |||||
Db 163 GAAAGGGCGTTGGTACATATGACAGGACACTCTTGACCGTTGGCATACGTTGCAAGA 222
    |||||
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAsp-IleLeuValGluAs 60
    |||||
Db 223 GCGTTGGCGGCAAAATCAGCTGAAGAAAGTAAGCGACACTACGAGCTCTCATTAGGGA 282
    |||||
QY 60 pLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThr-PheGluSera 80
    |||||
Db 283 TGTCAATGACATTGAGTCAGGCGCTTATCCACATCCCAATTACCGTTCAAATGGAACAA 342
    |||||
QY 80 snSerArgSerIle-AsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMet 99
    |||||
Db 343 CCACCTGAAGCATTAAGGAACCTCCAA--AAGGCTCTACTGAAGTGG----- 387
    |||||
QY 100 LeuSerIleTyrPheAspAenHisSerSerAspPheGluLys 113
    |||||
Db 388 -----ACATGGTTCCAAAACCAATTCACAAAGCACATCTAAG 423

RESULT 15
ADJ43786/C
ID ADJ43786 standard; cDNA; 466 BP.
XX
AC ADJ43786;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant cDNA #4786.
XX
KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
antifungal.
XX
OS Eukaryota.
XX
PN US2004016025-A1.
XX
PD 22-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00260238.
XX
PR 26-SEP-2001; 2001US-0325277P.
PR 26-SEP-2001; 2001US-0325448P.
PR 04-APR-2002; 2002US-0370620P.
XX
PA (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
```

PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICHKE D.
PA (ZHUT/) ZHU T.
XX
PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2004-190374/18.
XX
PT New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
PS Example 13; SEQ ID NO 4786; 230pp; English.
XX
CC The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 466 BP; 93 A; 127 C; 124 G; 122 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,38e-26 Length: 466
Score: 276.00 Matches: 66
Percent Similarity: 60.0% Conservative: 18
Best Local Similarity: 47.1% Mismatches: 27
Query Match: 42.1% Indels: 30
DB: 12 Gaps: 4

US-10-697-787-2 (1-126) x ADJ43786 (1-466)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProTrpThrPheSerGlnAsnLysMetPhe 20
Db 425 ATGGCAACAGCTCGGGAGCTCGTGC-----TGGACCGGAGCAGAACAGATGTTTC 372
QY 21 GluArgAlaLeuAlaValThrAspLysAspThrProAspArgTrpHisnValalLys 40
Db 371 GAGAAAGCCCTCGCGGTGTACGACGAGGACACGCCGACCGCTGGCACAACCGTGGCGGC 312
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 311 GCCGTGCGGGCAAGTACGAGGAGAGAGTGAAGCGGCACTACGAGCTGCTGGTGGCGGAC 252
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyr----- 74
Db 251 ATTATCCTCATCGAAGAGGCCAGATGCTCGAGGCAATTACCGCTCTCCGGCCACAGG 192
QY 75 -----LysThrPheGluSerAsnSerArgSerIleAsnAspPhe 87
Db 191 GGATGAAGTGTCTGATCAAAAGTGTAGAAACGGGACGAG-TCTACTCTGGATGAT 133

QY 88 AspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIleTyrPheAspAsnHis 107
Db 132 CAATTACAAGAAACAGATACATACATATATGTGATG----- 97
QY 108 SerSerAspPheGluLysPheSerGlnLysValLeu---ValSerTyrIleSerLeuVal 126
Db 96 -----AAGTATCAGACATCCCTGCTCTAAGTAGTGTCTGTTCTCTTGTTC 52

Search completed: March 10, 2006, 19:27:24
Job time : 480 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, **using frame plus p2n model**

Run on: March 10, 2006, 19:17:45 ; Search time 3944 Seconds

(without alignments)
1815.993 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNRSSISPTWFSQNMKFM.....HSSDFEKSQKVLVSIVSLV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB_spool/US10697787/runat_09032006_150627_3319/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absos8
-USER=US10697787 @CGN 1.1.4919 @runat_09032006_150627_3319 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.hcg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|--------------------|
| 1 | 656 | 100.0 | 381 | 15 | AY519525 Arabidops |
| 2 | 656 | 100.0 | 68041 | 15 | AC025814 Arabidops |
| 3 | 473 | 72.1 | 294 | 6 | CS138004 Sequence |

| | | | | | |
|----|-------|------|--------|----|---------------------|
| 4 | 344.5 | 52.5 | 334 | 15 | BT005657 Arabidops |
| 5 | 344.5 | 52.5 | 580 | 15 | AK119034 Arabidops |
| 6 | 343 | 52.3 | 303 | 15 | AY519524 Arabidops |
| 7 | 343 | 52.3 | 120977 | 15 | AC025808 Genomic s |
| 8 | 296 | 45.1 | 54433 | 15 | AC007119 Arabidops |
| 9 | 293 | 44.7 | 108844 | 15 | AC157502 Medicago |
| 10 | 289 | 44.1 | 246 | 15 | BT011255 Arabidops |
| 11 | 289 | 44.1 | 506 | 15 | BT010770 Arabidops |
| 12 | 286 | 43.6 | 306 | 15 | AY519526 Arabidops |
| 13 | 286 | 43.6 | 337 | 15 | BT008554 Arabidops |
| 14 | 286 | 43.6 | 570 | 15 | BT008698 Arabidops |
| 15 | 286 | 43.6 | 578 | 6 | CS137904 Sequence |
| 16 | 281 | 42.8 | 100108 | 14 | AP008111 Lotus cor |
| 17 | 276.5 | 42.1 | 89035 | 15 | AL050351 Arabidops |
| 18 | 276.5 | 42.1 | 198987 | 15 | AL161594 Arabidops |
| 19 | 276 | 42.1 | 3213 | 15 | AY954971 Antirrhin |
| 20 | 276 | 42.1 | 198354 | 15 | ATP22 |
| 21 | 276 | 42.1 | 198750 | 15 | ATP22 |
| 22 | 275.5 | 42.0 | 294 | 6 | CS137908 Sequence |
| 23 | 275.5 | 42.0 | 294 | 15 | AY519527 Arabidops |
| 24 | 274.5 | 41.8 | 3605 | 15 | AJ583670 Lycopersi |
| 25 | 263 | 40.1 | 612 | 15 | LES277944 Lycopersi |
| 26 | 261.5 | 39.9 | 97268 | 15 | AP004546 Lotus cor |
| 27 | 261.5 | 39.9 | 102550 | 14 | AP008039 Lotus cor |
| 28 | 261.5 | 39.9 | 108878 | 14 | AP007808 Lotus cor |
| 29 | 260 | 39.6 | 115071 | 14 | AC135566 Medicago |
| 30 | 254.5 | 38.8 | 166907 | 14 | AC146974 Zea mays |
| 31 | 252 | 38.4 | 512 | 15 | AK112054 Oryza sat |
| 32 | 252 | 38.4 | 34070 | 15 | AP007251 Oryza sat |
| 33 | 252 | 38.4 | 110000 | 15 | AP008207 254 |
| 34 | 252 | 38.4 | 110220 | 14 | AP003688 |
| 35 | 252 | 38.4 | 162297 | 15 | AP004317 Oryza sat |
| 36 | 251 | 38.3 | 91322 | 14 | AP007301 Lotus cor |
| 37 | 246.5 | 37.6 | 99587 | 15 | AC007858 Oryza sat |
| 38 | 246.5 | 37.6 | 110000 | 15 | AP008211 286 |
| 39 | 246.5 | 37.6 | 156643 | 15 | AC120988 Oryza sat |
| 40 | 245.5 | 37.4 | 92861 | 15 | AC006439 Arabidops |
| 41 | 245 | 37.3 | 513 | 6 | CS137972 Sequence |
| 42 | 244 | 37.2 | 72767 | 14 | AP007919 Lotus cor |
| 43 | 241 | 36.7 | 108564 | 14 | AC162787 Medicago |
| 44 | 218.5 | 33.3 | 110000 | 15 | AP008218 205 |
| 45 | 218.5 | 33.3 | 143618 | 15 | CNS08C7N |

ALIGNMENTS

| | | | | | |
|------------|------------|--|--------|--------|-----------------|
| RESULT 1 | AY519525 | Arabidopsis thaliana MYB transcription factor (Atig575250) mRNA, | 381 bp | linear | PLN 07-FEB-2004 |
| LOCUS | AY519525 | complete cds. | | | |
| DEFINITION | AY519525.1 | GI:41618977 | | | |
| ACCESSION | AY519525 | Arabidopsis thaliana (thale cress) | | | |
| VERSION | AY519525 | Arabidopsis thaliana | | | |
| KEYWORDS | | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi | | | |
| SOURCE | | Qu,L. and Gu,H. | | | |
| ORGANISM | | The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis | | | |
| REFERENCE | | Unpublished | | | |
| AUTHORS | | 2 (bases 1 to 381) | | | |
| TITLE | | Qu,L. and Gu,H. | | | |
| JOURNAL | | Direct Submission | | | |
| REFERENCE | | Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China | | | |
| AUTHORS | | Location/Qualifiers | | | |
| TITLE | | 1. 381 | | | |
| JOURNAL | | /organism="Arabidopsis thaliana" | | | |
| FEATURES | | source | | | |

gene /mol_type="mRNA"
1. .381 /db_xref="taxon:3702"
CDS /locus_tag="Atlg75250"
1. .381 /locus_tag="Atlg75250"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AA090995.1"
/db_xref="GI:41618978"
/translations="MASNSRSPSSPWFSSKNKFERALAVYDKTDPRHNVAKAVGG
KTVBEVKRHYDIIVELINIEGRVPLPNKYKTFESRSRINDPDTYIKLYLWMLSI
YFDNHSDFEKFSQKLVSYSLV"

ORIGIN

Alignment Scores:
Pred. No.: 5,92e-63 Length: 381
Score: 656.00 Matches: 126
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-697-787-2 (1-126) x AY519525 (1-381)

Qy 1 MetAlaSerAenSerArgSerSeriLeSerProTfPThrPheSerGlnAenLysMetPhe 20
Db 1 ATGCGGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATGCTC 60
Qy 21 GluAargAlaLeuAlaValTyAapLysAspThrProAspArgTrpHisAanValAlaLys 40
Db 61 GAGAGGCGCCTTGGCAGTTTACGACAGGACACACCCGACGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyAepLleLeuValGluAap 60
Db 121 GCTGTGCGAGGAGAACTGTAGAGAGTAGAGGCGCACCTATGACATCTCGTCGAGGAT 180
Qy 61 LeuIleAenTleGluThrGlyArgValProLeuProAenTyriLysThrPheGluSerAen 80
Db 181 CTCATCAACATCGAGACGTGGTGTGCTCCCTTGGCCCAATTACAAGACCTTCGAATCTTAAC 240
Qy 81 SerArgSerIleAenAapPheAapThrArgTyriLleThrLysTyriLeuTyriMetLeu 100
Db 241 TCAAGAGACATCAATGACTTTGACACAGGATGATATACTAAATATCTATATGATGCTC 300
Qy 101 SerIleTyriPheAapAenHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATATCATCTAGTGATTTTGAGAAATCTCTCAAAAAGTCTTGTA 360
Qy 121 SerTyriLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378

RESULT 2
AC025814 68041 bp DNA linear PLN 22-JAN-2001
LOCUS Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence,
complete sequence.
AC025814
AC025814.7 GI:12331602
VERSION HTG.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J.,
Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R.,
Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence
Unpublished
TITLE Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence
JOURNAL
REFERENCE 2 (bases 1 to 68041)

AUTHORS
TITLE Direct Submission
JOURNAL
REFERENCE 3 (bases 1 to 68041)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL
REFERENCE 4 (bases 1 to 68041)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL
REFERENCE 5 (bases 1 to 68041)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL
REFERENCE 6 (bases 1 to 68041)
AUTHORS Town,C.D. PhD.
TITLE Direct Submission
JOURNAL
COMMENT On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to:at@tigr.org

Lin,X. and Kaul,S.
Direct Submission
Submitted (15-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 68041)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (12-SEP-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
4 (bases 1 to 68041)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (05-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
5 (bases 1 to 68041)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (12-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
6 (bases 1 to 68041)
Town,C.D. PhD.
Direct Submission
Submitted (22-JAN-2001)
On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to:at@tigr.org

BAC clone F22H5 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Gensecan+ (Chris Burge
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant
of Glimmer, see Mihaela Pertea,
http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and
GeneSplicer (Mihaela Pertea and Steven Salzberg, contact
mpertea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

FEATURES
source
1. .68041
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosomes="1"
/clone="F22H5"
/ecotype="Columbia"
complement(418..5404)
/genes="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/genes="F22H5.15"
/codon_start=1
/product="unknown protein, 3' partial; 5120-134"
/protein_id="AA012690.1"
/db_xref="GI:10092277"
/translation="MEYEKSPATTTSRKISNNRSHLSFSAANAVDGVSSPVSNSKS
PLVDYGEI FRGSGPSSIPFLDVLNVLGVKVDVRSSKLDYSSVFGGACDFATV
PKEVI IKSCKTSINEDKRRNRKGGNSDVLPLCNKGKSPKPMVRMKHSDISYHQTV
RNENGATHLTQVPATPGPIPTQVDVNTSLHLKIESKSTPIPAVEKKLPCNREGEVK


```

|||||...:|||||...|||||...|||||...|||||...|||||...|||||...
Db 1 ATGGCTCTAGTCTTCTATGAGCTCGAGCTCTTCTTGGACGCTTAAGCAAAACAAGATGTTTC 60
QY 21 GUAAGAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisenValAlaLys 40
Db 61 GAGAGGGCTTTACCGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGCGGTAA 120
QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyAspLysLeuValGluAsp 60
Db 121 GCAGTTGGAGTAATCTGCAGAGGAAGTTAAAGCTCACTACGACATCCTCGTTGAAGAT 180
QY 61 LeuLeuAsnIleGluThrGlyArgValProLeuProAsnTyLysThrPheGlu----- 78
Db 181 CTCATGAACATCGAACCAAGACTTAGTACCTTTCCCTAAATACAAAACCGTCGATGTGGA 240
QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyLysThrLysTySyr----- 95
Db 241 AGTAATCTAGAGGCATCGATGTTTGGTGAAGTAAATGAGAAATATGAGAAATCCAG 300
QY 96 -----LeuTyMetMet 99
Db 301 TGAAGTTGTTACATGATG 318

```

RESULT 5
AK119034
LOCUS
DEFINITION
Arabidopsis thaliana At1g19510 mRNA for putative myb-related protein, complete cds, clone: RAF121-36-F10.
AK119034
ACCESSION
VERSION
GI:26453067
SOURCE
FLI CDNA; CAP trapper.
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Published Only in Database (2002)
2 (bases 1 to 580)
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: mseki@gsc.riken.go.jp, URL: http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625, Fax: 81-45-503-9586)

COMMENT
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector.
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
source
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF121-36-F10"
/ecotype="Columbia"
/note="common name: thale cress"
1..580
/gene="At1g19510/F18014_31"
88..390
/gene="At1g19510/F18014_31"

```

/codon_start=1
/product="putative myb-related protein"
/protein_id="BAC43610.1"
/db_xref="GI:26453068"
/translation="MASSMSSSSSWTSKONKMFERALAVYDKTDPDRWONVAKAYGS
KSABEVKGRHYDILVEDLMNIEQDLVPLPKYKTVGVGSKSRGIDDFDLRLMKNNRIQ"

```

ORIGIN

Alignment Scores:
Pred. No.: 1.58e-28 Length: 580
Score: 344.50 Matches: 72
Percent Similarity: 73.6% Conservative: 6
Best Local Similarity: 67.9% Mismatches: 21
Query Match: 52.5% Indels: 7
DB: 15 Gaps: 2

US-10-697-787-2 (1-126) x AK119034 (1-580)

QY 1 MetaLaserAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 88 ATGGCTCTAGTCTTCTATGAGCTCGAGCTCTTCTTGGACGCTTAAGCAAAACAAGATGTTTC 147

QY 21 GUAAGAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisenValAlaLys 40
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 148 GAGAGGGCTTTAGCGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGCGGTAA 207

QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyAspLysLeuValGluAsp 60
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 208 GCAGTTGGAAGTAATCTGCAGAGGAAGTTAAACGCTCACTAGCAGCATCCTCGTTGAAGAT 267

QY 61 LeuLeuAsnIleGluThrGlyArgValProLeuProAsnTyLysThrPheGlu----- 78
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 268 CTCATGAACATCGAACCAAGACTTAGTACCTTTCCCTAAATACAAAACCGTCGATGTGGA 327

QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyLysThrLysTySyr----- 95
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 328 AGTAAATCTAGAGGCATCGATGTTTGGTGAAGTAAATGAGAAATATGAGAAATCCAG 387

QY 96 -----LeuTyMetMet 99
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 388 TGAAGTTGTTACATGATG 405

RESULT 6
AY519524
LOCUS
DEFINITION
Arabidopsis thaliana MYB transcription factor (At1g19510) mRNA, complete cds.
ACCESSION
AY519524
VERSION
AY519524.1 GI:41618973
KEYWORDS
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
1 (bases 1 to 303)
Qu, L. and Gu, H.
AUTHORS
The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis
TITLE
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 303)
Qu, L. and Gu, H.
AUTHORS
Direct Submission
JOURNAL
TITLE
Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China
FEATURES
source
1..303
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..303
/locus_tag="At1g19510"

gene
1..303
/locus_tag="At1g19510"

| | | | |
|---|--|----------------------|---|
| CDS | 1. .303 /locus tag="At1g19510" /codon_start=1 /product="MYB transcription factor" /protein_id="AA099994.1" /db_xref="GI:41618974" /translations="MASSNSSSSSWTSQNKFFERALAVYKDTDPDRQNVAKVAGS KSAEVRKRYHILVEDLMNTEODLPKPKYKTVGVGSKRGIDDFDLRLMKNRRIQ" | JOURNAL | Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA |
| ORIGIN | | REFERENCE AUTHORS | 4 (bases 1 to 120977) Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, B., Chin, C., Chiou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J., Liu, S., Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vayenberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J. |
| Alignment Scores: | | TITLE | Direct Submission |
| Pred. No.: | 1.05e-28 | JOURNAL | Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA |
| Score: | 343.00 | REFERENCE | 5 (bases 1 to 120977) |
| Percent Similarity: | 84.0% | AUTHORS | Chao, Q., Brooke, S., Buehler, E., Johnson-Hopson, C., Khan, S., Kim, C., Shinn, P., Altafi, H., Bei, B., Chin, C., Chiou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vayenberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J. |
| Best Local Similarity: | 75.0% | | Direct Submission |
| Query Match: | 52.3% | | Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA |
| DB: | 15 | | On Apr 22, 2000 this sequence version replaced gi:7543634. |
| US-10-697-787-2 (1-126) x AY519524 (1-303) | | FEATURES | Location/Qualifiers |
| Qy 1 MetAlaSerAsnSerArgSerIleSerProTrpThrPheSerGlnAsnLysMetPhe 20 | | 1. .120977 | /organism="Arabidopsis thaliana" |
| Db 1 ATGGCTCTAGTTCTATGAGCTCGAGCTCTTCTTGGAGGCTTAAGCAACAAGATGTTTC 60 | | | /mol_type="genomic DNA" |
| Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAsnValAlaLys 40 | | | /db_xref="taxon:3702" |
| Db 61 GAGAGGGCTTTAGCGGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGTTGGCTAAA 120 | | | /chromosome="1" |
| Qy 41 AlaValGlyGlyIleValGluValGluValIleValArgHisTyAspIleLeuValGluAsp 60 | | 90. .2159 | /clone="F18014" |
| Db 121 GCATTTGGAAGTAATCTCGAGAGGAAGTTAAAGCTCCTACGACATCTCTGTTGAAGAT 180 | | | /notes="hypothetical protein" |
| Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyIleYsThrPheGlu----- 78 | | | /codon_start=1 |
| Db 181 CTCATGACATCGACACAGACTTAGTACCTTTCCTAAATACAAACCGTCGATGTTGGA 240 | | | /evidence="not experimental" |
| Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArg 90 | | | /product="F18014.1" |
| Db 241 AGTAAATCTAGAGGCATCGATGATTTTGTATTGAGG 276 | | | /protein_id="AAF79419.1" |
| RESULT 7 | | | /db_xref="GI:8778411" |
| AC025808/c | | | /translational="MISFEVSDVFTCSIVNAYCRSGNVDKAMVFAKETSGLGLEL NVYTNLSNGIAMIGDMGVMTGVRVRLMSRGVSRNVYVTSLSIKGYCKKGLMEAREH VFELKEKLVADQHMVMDGVCRTGQIRDAVRVHDNMIEIGVTRNTTICNSLING YCKSQLVEAEQIFSRMNDWSLKPDDHTYNTLVGCRAGYVDEALKLCCQMCQKEV PTVMTYNILLKGVSRICAFHDLVLSLWMLKRGNADEISCTSLLEALFKLGFNEAM KLWENVLARGLLTDITLVNMSGLCKMEKNEAKELIDNVNIFRCKPAVQTYQALSH GYKCNLKEAFVKEYMERKGFPTIEMVNTLISGAFKYRHLNKVADLVIELRARGL TPVTATYGALITGCNIGMIDKAYATCFEMIEKGIITLVNUNICSKIANSLRDLKIDEA CULLQKIVDFDLLPGYQSLKEPLEASATCTCKTKIAESVENSTPKKLLVPPNVYN VAIAGCKAGKLEDAKLFSDLLSDRFIPDEYTYTILHNGCAIGADINKATLRDEM ALKGITPNITVYNTALIKGLKLVNDVRAQRLHLKPKGITPNATYNTNTLIDGLVKSG NVAEAMRLKEKMIKGLVGRSKQGDVDIPKEVVLDPVEVKLGSTGVTEMNSNELYDVR RVSEAVV" |
| LOCUS | 120977 bp | | complement (4647. .5702) |
| DEFINITION | Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome 1, complete sequence. | | /notes="putative glycosyltransferase-like protein emb CAB42905.1; similar to EST gb A1998490.1" |
| ACCESSION | AC025808 | | /codon_start=1 |
| VERSION | AC025808.8 | | /evidence="not experimental" |
| KEYWORDS | HTG. | | /product="F18014.2" |
| SOURCE | Arabidopsis thaliana (thale cress) | | /protein_id="AAF79456.1" |
| ORGANISM | Arabidopsis thaliana | | /db_xref="GI:8778448" |
| REFERENCE | | | /translational="MSQHLILLILLLLHLPISATITTIQKPEAPQFVNSADCPPLI DSESDDDVVAKPIFCRRRAVHVMTLDAAYIRGSVAALVSLVLOHSCPENIVPHFVA SASADASSLRATISSSPYLDFTVYVFNVSIVSSIRSSISALDCPLNARSYLAOL LPPCVRVVYLDLILDLIAKLAATDLGRDSVLAAPCYCNANFTYELGSLRPPFLV LSLTPADRACVNTGVMDIQLSRREGAVTSIEEWAMQKMRIRIYELGSLRPPFLV FAGLIRPVNHRNQHGIGDGNFRGLCRDLHPGPVSLHLLHSGKGPWRLDAGRPCPLD ALWAPYDLLQTPFALDS" |
| AUTHORS | Shinn, P., Brooke, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, Q., Chin, C., Chiou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vayenberg, M., Yu, G., Federspiel, N.A., Theologis, A. and Ecker, J.R. | | complement (9620. .10300) |
| TITLE | Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome 1 | | /notes="similar to Zn finger protein gi 6319911" |
| JOURNAL | Unpublished | | /codon_start=1 |
| REFERENCE | 2 (bases 1 to 120977) | | |
| AUTHORS | Ecker, J.R. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA | | |
| REFERENCE | 3 (bases 1 to 120977) | | |
| AUTHORS | Ecker, J.R. | | |
| TITLE | Direct Submission | | |

```
/evidence=not experimental
/product="F18014.3"
/protein_id="AAF79455.1"
/db_xref="GI:8778447"
/translation="MSDVPSCSGNDTNNDSNFECNICLDLADQPIVTLGHLFCW
PCLYKWLHLSQKDCPCVKAVIDEDRLVPLXGRKSSADPRSKSIPLGLEVPNPSGQ
RPTAQPPDNGFAHHHGFMGFGAAPASARFGNVTLSAARFGLIPSLFNLHFH
GFPDAMYGAAASGGPFGHGSNPFHGGSHMSYQRHTGRQCGQDHLILLIVFVV
VVSLSFLS"
CDS
join(12523..12577,12686..13374)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.4"
/protein_id="AAF79420.1"
/db_xref="GI:8778412"
/translation="WAIFSTSHLLFISFIITACTTISVSGTTTTLNHCGSTIWPILIT
ANGAGDGGFALASGVSTFTVSPGWSGRFWARTYCNFDASGGKCGTDCGSKLKC
AGAGAPATLAEFTIGSGKNVADDFVSLVDGVYVQMKITPQSGSGDCKTAGCV
SDVNAICPKELQVTPSGVAACKSACEAFNKPKEYCCTGAYSTPATCPTPTNYSKIFQA
CPSAYSAYDDASSNFTCTNANYEISFCS"
CDS
complement(join(14020..14136,14236..14304,14402..14464,
14557..14696,14879..15005,16274..16486))
/note="unknown protein; similar to EST gb|H76493.1"
/codon_start=1
/evidence=not experimental
/product="F18014.5"
/protein_id="AAF79454.1"
/db_xref="GI:8778446"
/translation="MLEAVDSSGVVNGGFPQIQSYFGDCSSBEEISVLPRHTKVVVTG
NNRTKSVLQGVVKAQVGLGGMHLVLTNGIEVKLQNALSVLEPTTGNBDDDDLD
PNTQRNQSDEDTLKPXKSLRGQRSSRSHKTMSSLSLSDSSQSGSFGTPTPENM
KVDSLKEMPALLNRYHFNLDVAINPNSKQLIIVQHFMSQMDQLVIVGFVQA
AXRMKACFKSQKSRNTDLNCIS"
CDS
join(17994..19489,18568..18698,18814..18847,19006..19099,
19170..19347,19441..19479,19616..19715,19855..19994)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.6"
/protein_id="AAF79421.1"
/db_xref="GI:8778413"
/translation="MAKTDKLAQFLDSGIYESDEFNWFPLDTVTRNRSYTRFKVSPS
AYISRFNNSKOLNHSSESNNPKRRKOKNSFHLPSVGEQASNLRHQEARLFLSKAH
ESFLKEILLSITGLSDDDDDSSLLNKCCDDDEVSVFELGGVQWQAFYITILSPNL
HCDNGESCNRQVQPVNVLVNEIGEEVAFESNRRYIMPRNSCFYMSDLHHRNL
VPKSBEGNLIIVDPWENASAHQSKFVLYMYPTNPQYFLSLFKQLAHAFGAL
VALMYNREKLLSFVEKELFPWAGIKYVATMYWLKVPDGLICDLDLLAGSEKRSDF
KLLDKNQIMSPIDGFSRKPPIGDILLKHTPGSQPARCLELPAREMAAGWTSWGNEPL
HFQDSRYFLKY"
CDS
join(22037..22300,22618..23361)
/note="unknown protein; similar to ESTs gb|AI992723.1,
gb|T41777.1, dbj|AV440980.1, gb|AI997341.1, and
dbj|AV440785.1"
/codon_start=1
/evidence=not experimental
/product="F18014.7"
/protein_id="AAF79422.1"
/db_xref="GI:8778414"
/translation="MTSDGATSTSAALAAAAAATMRKPSWRENNRRRRRRRAVAA
KIYTLRAQGNVNLPHKCDNNELVKALCEAGVVEEDGTYTRKHGKPLPGDMAGSSS
RATPVSSHNQSPLASTEDPSILSYOVSPSSSPSPSPGDPHNIISTIFPLRNGIIP
SSLPPLRLSNAPVTPPVSSPTSRNPKPLPTWESFTKQSMNAKQSWTSNLYFYVAV
SAPASPTHRRQHPATIPCEDESSTVDSQHWISFOKFAQQQFPFSASWVPTFTFN
LVKPAPOQLSPNTAAIQIGSSSEFKFNSQVQKPEGERIHDVAMEDLELTGLNGKAH
S"
complement(join(23868..24983,25237..25407))
/note="unknown protein; similar to EST gb|T13858.1,
emb|Z38050.1, and emb|Z38087.1"
/codon_start=1
/evidence=not experimental
/product="F18014.8"
/protein_id="AAF79453.1"
```

```
/db_xref="GI:8778445"
/translation="MAGRRDRSQQLRSGRIATAILIGIFIGCVCAVLFPYGPNSSSS
LKASHELSKSNQVSSCSESRVAMLKSDFTLSEKNAELKKQVRETEKRLAEQ
GSDNARKQVLAGTQIKAGPFGTVKSLRTNPILPDESINPRLEIAVDKEVIV
ALANANVAMLEQVQIASIKRIGITNYLVALDDYIENLCKENDVAYTKRDPKQDVTV
KTKTGNHVASGLKFRVLREFQLQGVLLSDVDIVFLQNPFSHLVSDSVMSDGHG
NHTAYGFNDVDEPAMGWARVAHTNRIWVFNSEFFVLRPTIPISELLDRAVDRLSKAK
VWDOAVFNEELFYPSHPEVTHLASKRVMDMYEFNWSKVLFTVTRKHELKKVKPVI
VHVYTHPKLNRMQAVVEFYVNGKQDALDSFPDGSE"
complement(26100..27917)
/note="unknown protein; similar to ESTs gb|T42717.1,
emb|F13886.1, and emb|F13885.1"
/codon_start=1
/evidence=not experimental
/product="F18014.9"
/protein_id="AAF79452.1"
/db_xref="GI:8778444"
/translation="MADPLNGKFFICFSLFSFTLLFISPLYATESPVIEDVSTDA
VSVSTNREAVLLHNLNLTNLANLDAKLSATPLKBNBTSVDDDIQEEKERG
RAKASFVTKYSPFMSERFQTSVAKLNSDATCINVLPRDFEGSKYFAIGKGRVY
VFLRNGDVLIEFFTVDSFVTAMVSYSVFKNSSVFTVGHQNGAVLLHRIHGSNGED
WNSNSVMEHVKGKDDVDSADPVTLLVHHVGRVRYILATLSLGLTVLTENRTVYGS
VIPSRPLVFLKQRLFLFTESGAGSLDLRSKMKIRETECEGLNHSIARTYVFDAAERSK"
Alignment Scores:
Pred. No.: 1,48e-25 Length: 120977
Score: 343.00 Matches: 69
Percent Similarity: 80.4% Conservative: 5
Best Local Similarity: 75.0% Mismatches: 16
Query Match: 52.3% Indels: 2
DB: 15 Gaps: 1
US-10-697-787-2 (1-126) x AC025808 (1-120977)
Qy 1 MetAlaSerAsnSerArgSerSerIleSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 90486 ATGGCCTCTAGTCTTATGAGCTCGAGCTCTTCTTGACGCTCTAAGCAAAACGATGTT 90427
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPheHisnValalaLys 40
Db 90426 GAGAGGGCTTTAGCCGTTTACGATAAAGACATCCCGACCGTTGGCAAAACGTGCTAAA 90367
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 90366 CGAGTTGGAAGTAATCTCGCAGGAGAGTAACTAACCTACTACGACATCTCTGTAAGAT 90307
Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu----- 78
Db 90306 CTCATGAACATCGAACAGACTTAGTACCTTTGCTTAATAACAAACCGTCGATGTGGA 90247
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArg 90
Db 90246 AGTAAATCTAGAGCATCGATGATTTGATTGAGG 90211
RESULT 8
AC007119/7 AC007119 54433 bp DNA linear PLN 27-FEB-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F2G1 map m1238, complete
DEFINITION sequence.
AC007119
VERSION AC007119.6 GI:20198094
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 54433)
Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,
Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
Nierman,W.C., Fraser,C.M. and Venter,J.C.
JOURNAL Unpublished
```

| | | | |
|--------------|---|---------------|--|
| REFERENCE | 2 (bases 1 to 54433) | | /db_xref="GI:20198095" |
| AUTHORS | Lin.X. | | /translation="NATNVEYTCFVRGLDQDTDEKDLTDIFSKFGNVIDSKIIYDRDT |
| TITLE | Direct Submission | | GRSRFGFVTFEEKSMTDALIMDVESRSKCVNNGSVITVEARQRKRRAEFALE |
| JOURNAL | Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA | gene | LVRLLNEKQKQKAGL" |
| REFERENCE | 3 (bases 1 to 54433) | | /complement (6969. .8258) |
| AUTHORS | Town,C.D. and Kaul,S. | | /gene="At2g21680" |
| TITLE | Direct Submission | | /note="synonym: F2G1.5; predicted by genefinder" |
| JOURNAL | Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org | mRNA | complement (join(<6969. .7523,7569. .>8258)) |
| COMMENT | On Apr 18, 2002 this sequence version replaced gi:6598668. Sequencing, analysis, and annotation were performed within the CSHL/WUGSC/ABI Arabidopsis Genome Sequencing Consortium. Information on physical mapping and YAC and BAC library construction as well as added annotation can be viewed at http://www.cshl.org/arabweb/ . We used GenScan, Grail, and MZEF for predicting coding exons and assembling genes. BAC F6P23 maps to YAC | CDS | /complement (join(6969. .7523,7569. .8258)) |
| FEATURES | YUPCIC9B1. | | /gene="At2g21680" |
| source | Location/Qualifiers | repeat_region | /codon_start=1 |
| | 1. .54433 | gene | /product="hypothetical protein" |
| | /organism="Arabidopsis thaliana" | | /protein_id="AAD23638.1" |
| | /mol_type="genomic DNA" | | /db_xref="GI:4567223" |
| | /db_xref="taxon:3702" | tRNA | /translation="WVLISEYDDSGTGGDHOIKPKPEEDRNKKLKEKQVSLPIPE |
| | /map="mi238" | | ELILRCFLVRRCHHPISLSVCRISDFSLMSKLYDDRLRLGYTENLYAVGPPVPNP |
| | /clone="F2G1" | | SWYILHRPYRNLPTNISLKCISDLPMPMGSTVVTIGSDIYVIGGVRGKLLGDV |
| | /ecotype="Columbia" | | GVGNKPISSGREGTSIRGGHAGERRISDVTHINCRPHEYRSLPSMKMARCPAAAGV |
| misc_feature | complement (1. .6910) | | IDGKIYVIGGRKSSSVPGPYPEAFGRGEFLTAYVMKEKIYCLDLTRNTHIYDPKESK |
| misc_feature | /note="overlap with BAC clone F7D8 (AC007019:1. .6910)." | | WESWTHGPLSASWNSDSCVVDNLLFCINTSVIFLGWPIKIYDPEKKTWYLGQLQGF |
| | 1 | | ANGLFVDGYKMAFCGKLVLSADVHRLRRYDCRKREIWCIEIAMERKEDGTFWGKVE |
| | /notes="33956 bases before this point were not included in the submitted sequence, due to overlap with another BAC" | | SVAVLTPAKTTSDICGTIV" |
| | complement (2681. .5102) | gene | complement (8345. .8400) |
| | /note="synonym: F2G1.2; predicted by genscan" | | /rpt_family="AT_rich" |
| | complement (join(<2681. .2942,3015. .>3289,3419. .>3523,3592. .4315,4447. .>5102)) | | 9099. .9179 |
| | /gene="At2g21710" | | /gene="At2g21670" |
| | complement (join(2681. .2942,3015. .>3289,3419. .>3523,3592. .4315,4447. .>5102)) | | /product="tRNA-Leu" |
| | /gene="At2g21710" | gene | 9672. .10691 |
| | /codon_start=1 | | /gene="At2g21660" |
| | /product="hypothetical protein" | | /note="synonym: F2G1.4; Alternative splicing exists based on EST evidence, form 1 splice site is AG, form 2 splice site is AT" |
| | /protein_id="AA015397.1" | mRNA | join(<9672. .9785,10095. .>10691) |
| | /db_xref="GI:20198096" | | /gene="At2g21660" |
| | /translation="MLLHCNVSYTTFGFISSLRQDNADDSODTVIRRHARS1 | CDS | join(9672. .9785,10095. .>10511) |
| | SLYIRNRDLKKNPNESQETVPPPPRRDLDCGNSKLELSLTVRRTPQPPGS | | /gene="At2g21660" |
| | IYAQASDADIASLSLRFNFGSDGDDGSESRMIVKALEIRRKVKTEILKESLVR | | /codon_start=1 |
| | KGRFGITYATNVTDRGLDFVDHVMIAQAALKRUPFESGRFNLRARTVIEDSNFVLV | repeat_region | /product="glycine-rich RNA binding protein 7" |
| | RWLKHELSTNRIAKIICMSKGLDSIRIMIEWKSIHVKGESIAVAFIRSGDNILRF | | /protein_id="AAD23639.1" |
| | NREELNEIVLSNGVRDRMGVYVGRCPPELLSFSGMEVKSRVDFFLKMGMNQDFG | | /db_xref="GI:4567224" |
| | TMVYDPIKTIQFSFOVMEKKVLPALFNTPALRLSPKFTIIVLLVLAQINVLKPFGLST | | /translation="WASGDVEYRCFVGGLAWATDDPALETAFAQYGDVIDSKINDRE |
| | EVVGRLLAYPHLMGCSLEERWKPLVYFYVGLIPKEGMKRIIVKVPILYCIDLEXTI | | TGRSGFGVTFDKRAMKDAIEGMNQDLDRSITVNEAQRSGSGGGHRRGGGG |
| | APKVELRYNRFLOEMGIPNEAIGNMLVKFPSSLTNSLYKKIRPVIPLTRAGVTQK | | YRSGGGGYSGGGSGYSGGGRRGGGYSGGGGYSGRGGGSGYGGGRGGGGTG |
| | DIGKLTAMDALLGSGIGTKLEPNMRYISLGIRFYOLGEMIADFPMLRYNVNLRP | repeat_region | complement (9910. .9934) |
| | KYRLATMRIPQDIETFPFYSYLERRIIPRTHIVENRYNFKRLYLACTDEF | | /rpt_family=" (TAG)n" |
| | ERRRVKVERFERAGLSDSQSDNISQDIAFSADEAEEDITE" | repeat_region | 10224. .10311 |
| | 5254. .5325 | | /encoding glycine-rich proteins e.g., CCR2, L04172; GRP7, Z14987, 20 copies#Rf:ATR0048 X93610 182A" |
| | /gene="At2g21700" | | 10246. .10390 |
| | 5254. .5325 | | /rpt_family=" (CGG)n" |
| | /product="tRNA-Gln" | repeat_region | 10400. .10505 |
| | 5391. .5874 | | /rpt_family=" (GGA)n" |
| | /gene="At2g21690" | gene | complement (11890. .12451) |
| | /note="synonym: F2G1.1" | | /genes="At2g21655" |
| | join(<5391. .5501,5632. .>5874) | mRNA | /notes="synonym: F2G1.3; supported by full length cDNA: |
| | /gene="At2g21690" | | Cere8: 10836" |
| | /codon_start=1 | | complement (<11890. .>12451) |
| | /product="putative glycine-rich RNA binding protein" | | /gene="At2g21655" |
| | /protein_id="AA015396.1" | CDS | complement (11934. .12404) |
| | | | /gene="At2g21655" |
| | | | /codon_start=1 |
| | | | /product="Expressed protein" |
| | | | /protein_id="AA015398.1" |
| | | | /db_xref="GI:20198097" |
| | | | /translation="MENKTMFIMFSLIMVLLSF5SHPTFGKESNDKPLLLSDSDFDAM |
| | | | MARSPTSNDYNENVGSKQINLMKNSKQMAVPDKCIEEYMAEIQNKASRDCC |
| | | | LGTIVKAGKECHMEYMKLFFQMYELRRFTSKRFSKTNEIWNRCSTEIGVWSPYSG |
| | | | complement (13202. .13225) |

```

repeat_region /rpt_family="AT rich"
complement(14063..14085)
/rpt_family="AT rich"
gene complement(15427..16406)
/gene="At2g21650"
/note="synonym: F2G1.8"
mRNA complement(join<15427..15593,16054..>16406))
/gene="At2g21650"
CDS complement(join(15569..15593,16054..16334))
/gene="At2g21650"
/codon_start=1
/product="unknown protein"
/protein_id="AAD23640.1"
/db_xref="GI:4567225"
/translation="MASGSMSSYSGSWTVKONKAFERALAVYDQDTPRWNVARAV
GGKTPPEAKRQYDLVRDIESLENGHVPFDYKTTTGNNSRGRLDRBEKMRSMKIQ"
complement(19805..19831)
repeat_region /rpt_family="(A)n"
complement(20361..21003)
/gene="At2g21640"
/note="synonym: F2G1.9"
mRNA complement(join<20361..20639,20794..>21003))
/gene="At2g21640"
CDS complement(join(20535..20639,20794..21003))
/gene="At2g21640"
/codon_start=1
/product="unknown protein"
/protein_id="AAD23641.1"
/db_xref="GI:4567226"
/translation="MASQNTQVTWTKIETENPPKQVPASSCRKRVKDDNATPFAN
LKHDMDEFTHASMDHKTCFKNTMDKIFGFSKAEEAEOKEAEVVEIHSPLQTA
TK"
repeat_region complement(21486..21567)
/note="synonym: (GAAA)n"
gene complement(22159..25642)
/gene="At2g21630"
/note="synonym: F2G1.10; similar to human SEC23"

Alignment Scores:
Pred. No.: 8,18e-21 Length: 54433
Score: 266.00 Matches: 60
Percent Similarity: 70.8% Conservative: 8
Best Local Similarity: 62.5% Mismatches: 24
Query Match: 45.1% Indels: 4
DB: 15 Gaps: 2

US-10-697-787-2 (1-126) x AC007119 (1-54433)
QY 1 MetAlaSerAsnSerArgSerSerile-----SerProTrrPThrPheSerGlnAsnLys 18
Db ATGGCATCAGGCTCAATGCTCTTCTTATGGCTCTGGCTCATGCTGTTAAGCAGAACAAA 16275
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
Db GCCTTTGACGGTCTAGCAGTCTATGACCAAGACACTCCGACCGTGGCACAATGTT 16215
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Db GCTAGAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 16214
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTrrLysThrPheGlu 78
Db CGTGACATCGAAGCATCGAGAATGTCACGTCGTCATTCCTGCTACTACAGACTACTACA 16095
QY 79 SerAsnSerArg-----SerIleAsnAspPheAspThrArgTyrIle 92
Db GGAACAGACACAGACAGCGGCTGCTGTATGAGGAAAGAGGTATATT 16047

RESULT 9
AC157502 108844 bp DNA linear PLN 13-MAY-2005
LOCUS Medicago truncatula chromosome 7 BAC clone mte1-56e23, complete
DEFINITION sequence.

```

```

ACCESSION AC157502
VERSION AC157502.2 GI:63987073
KEYWORDS HTG
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 108844)
Town, C.D., Tallon, L.J., Arbogast, T., Althoff, R., Hine, E.,
Monaghan, E., Smith, S.A., Utterback, T., Feldblyum, T., Koo, H. and
Cheung, F.
Medicago truncatula BAC genomic sequence
Unpublished
2 (bases 1 to 108844)
Town, C.D.
Direct Submission
Submitted (18-FEB-2005) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 108844)
Town, C.D.
Direct Submission
Submitted (13-MAY-2005) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT On May 13, 2005 this sequence version replaced gi:59933338.
FEATURES
source
1..108844
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/chromosome="7"
/clone="mte1-56e23"

ORIGIN
Alignment Scores:
Pred. No.: 4,04e-20 Length: 108844
Score: 293.00 Matches: 61
Percent Similarity: 61.8% Conservative: 23
Best Local Similarity: 44.9% Mismatches: 24
Query Match: 44.7% Indels: 28
DB: 15 Gaps: 4

US-10-697-787-2 (1-126) x AC157502 (1-108844)
QY 1 MetAlaSerAsnSerArgSerSerileSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db ATGGCATCTAGCTCAATGCTGCTCTGGCTCATGAGTGTAAAGGAGAAATAAGCCTTT 84631
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db GAAAGAGCTTTAGCTGTTTATGACAAAGACACACAGATCGTTGGTACAAATGAGTCAT 84691
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db GCTGTTGGTGGCAAACTCTCGAGGAGTTAAGAACTATGAACTTCTGTTGAGGAT 84751
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTrrLysThrPheGluSerAsn 80
Db ATTAACATATTTAGCTAGGGAAGTTCCATTTCCAAATTTACAGAAAATT----- 84802
QY 81 SerArgSerIleAsnAspPheAspThrArgTyr-----IleThrLysTyrLeuTyr 97
Db TCAGTGTCTCATGAGGAGAAAGGTATCATTAATAAATAGCTTCATATTC 84856
QY 98 MetMetLeuSerIle-----
Db CCTACATGTTGTTGTTAAATAATTAAACAGAGCTTATGAAGCAAAAGTTAAATATAC 84916
QY 103 TyrPheAspAsnHisSer-----SerAspPheGluLys 113
Db TACTATACTCTAACTCAACTTTTTTTTTTTTATATCTGATTTTTCATANA 84964

```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS

1 (bases 1 to 306)
Qu, L. and Gu, H.

TITLE

The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis

JOURNAL

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 306)
Qu, L. and Gu, H.

TITLE

Direct Submission

JOURNAL

Submitted (07-JAN-2004) Life Sciences. National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

FEATURES

source

1. .306
/organism="Arabidopsis thaliana"

gene

1. .306
/db_xref="taxon:3702"

CDS

1. .306
/locus_tag="At2g21650"

1. .306
/locus_tag="At2g21650"

1. .306
/codon_start=1

1. .306
/product="MYB transcription factor"

1. .306
/protein_id="AA09996.1"

1. .306
/translation="MAGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAV

1. .306
GGKTPEAKRQYDLLVRDIESIENGHPFDDYKTTGNSNRGLRDEKRMKMLQ"

1. .306
/db_xref="GI:41618982"

1. .306
/db_xref="taxon:3702"

1. .306
/locus_tag="At2g21650"

1. .306
/locus_tag="At2g21650"

1. .306
/codon_start=1

1. .306
/product="MYB transcription factor"

1. .306
/protein_id="AA09996.1"

1. .306
/translation="MAGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAV

1. .306
GGKTPEAKRQYDLLVRDIESIENGHPFDDYKTTGNSNRGLRDEKRMKMLQ"

1. .306
/db_xref="GI:41618982"

1. .306
/db_xref="taxon:3702"

1. .306
/locus_tag="At2g21650"

1. .306
/locus_tag="At2g21650"

1. .306
/codon_start=1

1. .306
/product="MYB transcription factor"

1. .306
/protein_id="AA09996.1"

1. .306
/translation="MAGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAV

1. .306
GGKTPEAKRQYDLLVRDIESIENGHPFDDYKTTGNSNRGLRDEKRMKMLQ"

1. .306
/db_xref="GI:41618982"

1. .306
/db_xref="taxon:3702"

REFERENCE
AUTHORS

1 (bases 1 to 337)
Yamada, K., Dale, J.M., Huan, V.W., Onodera, C.S., Quach, H.L., Chen, H., Toriumi, M., Wong, C., Yu, G., Yuan, S., Carninci, P., Kawai, J., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 337)
Yamada, K., Dale, J.M., Huan, V.W., Onodera, C.S., Quach, H.L., Chen, H., Toriumi, M., Wong, C., Yu, G., Yuan, S., Carninci, P., Kawai, J., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

The Salk, Stanford, PGENC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Dale, J.M., Huan, V.W., Onodera, C.S., Quach, H.L., Toriumi, M., Wong, C., Yu, G., Yuan, S., Carninci, P., Kawai, J., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE

Arabidopsis Open Reading Frame (ORF) Clones

| | | | | |
|--|---|---|------|---|
| DB: | 15 | Gaps: | 1 | Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank. |
| US-10-697-787-2 (1-126) x BT008554 (1-337) | | | | |
| Qy | 1 | MetAlaSerAanSerArgSerSerIle-----SerProTrpThrPheSerGlnAanLys | 18 | |
| Db | 1 | ATGGCATCAGGCTCAATGCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGAACAAA | 60 | |
| Qy | 19 | MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTtrPheHisAanVal | 38 | |
| Db | 61 | GCCTTTGAGCGTGTCTAGCAGCTATGACCAAGACACTCCGGACCGTTGGCAATGTT | 120 | |
| Qy | 39 | AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal | 58 | |
| Db | 121 | GCTAGAGCTGTGTGTGTTAAACACCAAGAAAGCTTAAGAGACAGTATGACCTTCTAGTT | 180 | |
| Qy | 59 | GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGlu | 78 | |
| Db | 181 | CGTGACATCGAAAGCATCGAATGGTCAGTACCATTCCTGACTACAGACTACTACA | 240 | |
| Qy | 79 | SerAanSer 81 | | |
| Db | 241 | GGAACACAGC 249 | | |
| RESULT 14 | | | | |
| BT008698 | | | | |
| LOCUS | BT008698 | 570 bp | mRNA | linear |
| DEFINITION | Arabidopsis thaliana clone RAFL17-18-C23 (R50297) putative myb family transcription factor (At2g21650) mRNA, complete cds. | | | PLN 16-MAY-2003 |
| ACCESSION | BT008698 | | | |
| VERSION | BT008698.1 | GI:30794123 | | |
| KEYWORDS | FLI CDNA. | | | |
| SOURCE | Arabidopsis thaliana (thale cress) | | | |
| ORGANISM | Arabidopsis thaliana | | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis. | | | |
| AUTHORS | 1 (bases 1 to 570) | | | |
| Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. | | | | |
| Arabisopsis Full Length cDNA Clones | | | | |
| Unpublished | | | | |
| 2 (bases 1 to 570) | | | | |
| Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. | | | | |
| Direct Submission | | | | |
| Submitted (16-MAY-2003) | | | | |
| Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA | | | | |
| RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. | | | | |
| The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A. | | | | |
| Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs. | | | | |
| FEATURES | | | | |
| source | 1..570 | | | |
| Location/Qualifiers | | | | |
| 1..570 | | | | |
| /organism="Arabidopsis thaliana" | | | | |
| /mol_type="mRNA" | | | | |
| /db_xref="taxon:3702" | | | | |
| /chromosome="2" | | | | |
| /clone="RAFL17-18-C23 (R50297)" | | | | |
| /ecotype="Columbia" | | | | |
| /note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert." | | | | |
| 1..570 | | | | |
| /gene="At2g21650" | | | | |
| 1..8 | | | | |
| /gene="At2g21650" | | | | |
| 9..314 | | | | |
| /gene="At2g21650" | | | | |
| /codon_start=1 | | | | |
| /evidence=experimental | | | | |
| /product="putative myb family transcription factor" | | | | |
| /protein_id="AAP40504.1" | | | | |
| /db_xref="GI:30794124" | | | | |
| /translation="MAGSSMSYSGSWTVKONKAFERALAVVDDTPDRWHNVARAV GSKTPEAKRQYDLLVRDIESIENGHPFDYKTTTGNRNRLRDEEXMRSMKLG" | | | | |
| misc_difference 221 | | | | |
| /gene="At2g21650" | | | | |
| /note="compared to genomic sequence" | | | | |
| /replaces="g" | | | | |
| 315..570 | | | | |
| /gene="At2g21650" | | | | |
| misc_difference 554 | | | | |
| /gene="At2g21650" | | | | |
| /note="not present in genomic sequence" | | | | |
| ORIGIN | | | | |
| Alignment Scores: | | | | |
| Pred. No.: | 4.12e-22 | Length: | 570 | |
| Score: | 286.00 | Matches: | 56 | |
| Percent Similarity: | 74.7% | Conservative: | 6 | |
| Best Local Similarity: | 67.5% | Mismatches: | 19 | |
| Query Match: | 43.6% | Indels: | 2 | |
| DB: | 15 | Gaps: | 1 | |
| US-10-697-787-2 (1-126) x BT008698 (1-570) | | | | |
| Qy | 1 | MetAlaSerAanSerArgSerSerIle-----SerProTrpThrPheSerGlnAanLys | 18 | |
| Db | 9 | ATGGCATCAGGCTCAATGCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGAACAAA | 68 | |
| Qy | 19 | MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTtrPheHisAanVal | 38 | |
| Db | 69 | GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGGACCGTTGGCAATGTT | 128 | |
| Qy | 39 | AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal | 58 | |
| Db | 129 | GCTAGAGCTGTGTGTGTTAAACACCAAGAAAGCTTAAGAGACAGTATGACCTTCTAGTT | 188 | |
| Qy | 59 | GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGlu | 78 | |
| Db | 189 | CGTGACATCGAAGCATCGAATGGTCAGTACCATTCCTGACTACAGACTACTACA | 248 | |
| Qy | 79 | SerAanSer 81 | | |
| Db | 249 | GGAACACAGC 257 | | |
| RESULT 15 | | | | |
| CS137904 | | | | |
| LOCUS | CS137904 | 578 bp | DNA | linear |
| DEFINITION | Sequence 875 from Patent WO2005047516. | | | PAT 09-AUG-2005 |
| ACCESSION | CS137904 | | | |
| VERSION | CS137904.1 | GI:72058614 | | |
| KEYWORDS | | | | |

Search completed: March 10, 2006, 21:12:36
Job time : 3966 secs

THIS PAGE BLANK (USPTO)